OPEN ACCESS

University of BRISTOL

This electronic thesis or dissertation has been downloaded from Explore Bristol Research, http://research-information.bristol.ac.uk

Author:
Parker, Richard

Title:
Cognitive bias as an indicator of emotional states in animals

General rights
The copyright of this thesis rests with the author, unless otherwise identified in the body of the thesis, and no quotation from it or information derived from it may be published without proper acknowledgement. It is permitted to use and duplicate this work only for personal and non-commercial research, study or criticism/review. You must obtain prior written consent from the author for any other use. It is not permitted to supply the whole or part of this thesis to any other person or to post the same on any website or other online location without the prior written consent of the author.

Take down policy
Some pages of this thesis may have been removed for copyright restrictions prior to it having been deposited in Explore Bristol Research. However, if you have discovered material within the thesis that you believe is unlawful e.g. breaches copyright, (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please contact: open-access@bristol.ac.uk and include the following information in your message:

• Your contact details
• Bibliographic details for the item, including a URL
• An outline of the nature of the complaint

On receipt of your message the Open Access team will immediately investigate your claim, make an initial judgement of the validity of the claim, and withdraw the item in question from public view.

COGNITIVE BIAS AS AN INDICATOR OF EMOTIONAL STATE IN ANIMALS

RICHARD M A PARKER

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of Doctor of Philosophy in the Faculty of Medical and Veterinary Sciences.

Department of Clinical Veterinary Sciences

September 2008

Word Count: 79,367

ABSTRACT

An important objective of animal welfare science is the development of indicators of putative subjective emotional (affective) states. To this end, Paul et al (2005) have proposed devising animal-based measures sensitive to changes in certain cognitive processes found to be biased in characteristic ways across affective state in humans. This thesis investigates the application of this approach.

The first three experimental chapters examine judgements of ambiguous stimuli in rodents. In the first two of these studies, it was hypothesised that a treatment designed to induce a positive, or negative, change in affect would be associated with a higher, or lower, probability (respectively) of responding to ambiguous stimuli in a manner in keeping with a bias towards optimism; such biases, across affect, have been found in humans. These hypotheses were not supported, at least not in simple terms, with the results revealing counter-intuitive treatment effects, and variation in response accuracy and efficiency. In the last of these three experimental chapters, we applied a treatment designed to induce a change in food motivation. This altered rats' operant responses in a manner suggesting their behaviour was a least partly goal-directed, and also suggesting that the possibility of motivation-related confounds, when studying responses to ambiguity, was real.

The final experimental chapter investigated affect-related biases in the foraging behaviour of domestic chicks. We hypothesised that chicks undergoing a treatment designed to induce a negative change in affect would attack fewer red crumbs (a colour commonly associated with aposematism), and more green crumbs, than a control group. We found the opposite: i.e. the former treatment group attacked significantly more red crumbs. This curious finding was reconciled with reference to the functional architecture of the attentional processes implicated in foraging behaviour.

In the final chapter, the implications of these, and related, findings are discussed.

ACKNOWLEDGEMENTS

Firstly, I would like to thank my supervisors, Mike Mendl and Liz Paul, for giving me the opportunity to conduct this work. In the acknowledgements to her thesis, a past student of their's praised Mike's patience and forbearance, despite the fact, she suggests, he must of felt like hitting her with a baseball bat at times. During my PhD, I suspect Mike has felt like wielding a whole baseball stadium, but I'm happy to report he has refrained from doing so. Indeed, far from dodging stadia-sized blows, I have enjoyed excellent advice, support and encouragement from both my supervisors, and also from Oliver Burman, who has been kind and helpful beyond the call of duty.

In addition, I'd like to thank/blame Bill Browne, for introducing me to MLwiN, and for his statistical advice. From Newcastle, thanks to Candy Rowe for inviting me to work in her lab, and to her, John Skelhorn, and Christina Halpin, together with the animal care staff, for the helpful assistance they provided during my visit. Thanks also to another Toon-dwelling academic, Melissa Bateson, for useful discussions on theoretical and methodological matters.

Back in Langford, I'm grateful to all those who have contributed to the fun and supportive milieu in which I have worked: Poppy Statham, Sam Jones, Becky Hothersall, Mike Toscano, Sam Gaines, Luca Melotti, Usama Abou Ismail, Jo Edgar, Gemma Richards, Bex Shackleton (my 'blind' video analyst), Sarah Lambton, Mary Friel, Eimear Murphy, Cathy Trickett, Simon Hill, Christine Basse, Suzanne Held, Stuart Pope, Mike Kitcherside, Steve Brown, Roz Ellis, Zoe Demery, together, I suspect, with many others I have forgotten due to post-write-up fatigue.

In addition, I'm grateful to Sean Rands, Val Warris, and most recently, to Bill Browne, for keeping me in gainful employment over the past year.

Finally, I am very thankful for the unwavering support I have received from all my family, and friends outside academia, and to Nikki Rooney, for her huge encouragement, and strategic deployment of thesis-deadline-related gifts.

AUTHOR'S DECLARATION

I declare that the work in this dissertation was carried out in accordance with the Regulations of the University of Bristol. The work is original, except where indicated by special reference in the text, and no part of the dissertation has been submitted for any other academic award. Any views expressed in the dissertation are those of the author.

Signed: ..... Date: ..... 4/6/09 .

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

A NOTE FOR THE READER

Please note, that in keeping with recent trends in scientific prose, portions of this thesis are written in the first person. Furthermore, since I wrote this document with a view to adapting parts of it for co-authored publication, I often use the plural pronoun, 'we'.

In addition, I hope you are not too daunted by the length of this document. The appendix takes up a considerable portion of it, and this requires only cursory glances, at most. Furthermore, I have employed a number of footnotes; these are, of course, optional reading: asides, or minor clarifications, which are not central to the understanding of the main text, but which will hopefully be of interest to some. Finally, a few of the Results sections are rather in-depth; whilst that's not a bad thing, where they risk becoming too dense, however, I have placed summaries at the end of subsections, and have otherwise endeavoured to explain the techniques I have used as clearly as possible.

CHAPTER 1

GENERAL INTRODUCTION

PROLOGUE

You might think I'm writing this consciously: i.e. that I'm subjectively aware of my surroundings, the screen, the flow of words as I edit them in my 'mind's eye' before putting them on the page, but actually I'm not - I'm not consciously aware of any of those things, nor the fact that there's a dog barking in the distance, nor that my feet are a bit cold. In fact, I'm not conscious of anything: never have been, never will be. You can find out for yourselves: whilst I'd appreciate it if you didn't open my brain up, I'll happily let you see the results of a whole range of scans and tests I've recently undergone. To be honest, they all show the same thing: everything you need to know about how I work (how I type, think, react, speak, play, moan, smile, etc.) is right there (or here), in my body: i.e. it's all in the bag of molecules and matter which make up me. It's very clever, the way my body does it, and whilst scientists are still sketchy on some of the details, if you just open up a textbook on physiology or anatomy, maybe one on neuronal functioning, cognitive science, endocrinology, perhaps even a couple on basic physics, you'll find all the answers you need there; OK, maybe not *all* of them, but they'll at least give you a reasonable idea about what those sketchy details might turn out to be.

In fact, I *used* to think I was conscious, but then a scientist pointed out that I wasn't: she used a clockwork toy to show me – she opened it up, pointed at the wind-up mechanism, noted how it connected to the monkey's arms and legs, how it made them move, and the cymbals beat, how it transformed stored energy, how the cogs told the arms to move at a particular rate, how the gears co-ordinated all the actions, how a bit of oil helped it along from time-to-time, and how the whole thing stopped of its own accord; her explanation was certainly comprehensive – after she'd talked me through it, I was satisfied I knew how the toy worked.

Then she showed me the results of those scans and tests, and also those textbooks I was talking about: she explained how I sensed what was going on

around me, how that information was streamed, processed, integrated; how my body gained energy, then burnt it up; how and where my body made decisions; how I learnt, remembered, and forgot things; how I moved, acted, spoke; how I grew, adapted, degenerated; and also how some of these processes changed when the state of my system changed: e.g. when I was hungry (i.e. energy-depleted), in pain (maybe neurons were delivering messages from a toe I stubbed earlier), 'anxious' (perhaps I'd just seen a bear, or remembered the deadline for thesis submission), and so on, key aspects of my body changed in an adaptive way (I told you it was clever!)

As with the clockwork toy, her explanation of how *I* worked seemed pretty comprehensive, but of course she'd left out the 'conscious' bit: the part of me that was 'subjectively aware', the experiential 'I' that conducted all those mechanical instruments. Not so, she said: recall we found no such 'conscious' or 'subjective' element in the clockwork toy, nor did we need to cite it when explaining how it worked: we were quite happy and satisfied that what we saw was all there was to know about the toy, so why should the rules be changed for humans? Clearly, she didn't think I was conscious at all! I objected, of course, protesting that I really *was* 'aware', and my consciousness definitely *did* something – surely it would be there if we looked hard enough... but then she told me exactly *how* I was protesting (she got the books and charts out again...), and that if she had enough time and money, she could get the clockwork toy (or a sophisticated version of it) to protest that *it* was in fact conscious too: yet when we opened it up again, we'd still find just the mechanisms, nothing else, and we'd know exactly how they (and she) had done it.

And actually, when she put it like that, there didn't seem to be any room for 'me' at all: no role for the conscious, 'subjective' bit I used to think did all those things! It's a bit disconcerting at first, but when you look at the machine closely, the ghost just seems to disappear. So, anyway, that's how I became non-conscious. Just like a tree, a table, a rock, and that barking dog.

THE PROBLEM

Whilst these opening lines contain a conceit or two, they nevertheless illustrate some issues peculiar to the study of the mind, consciousness[1], subjective awareness, and emotion. Science is an objective endeavour; putting aside the problematic caveat that it is conducted by people who inevitably engage with science through the prism of their (apparently!) subjective experiences, it's nevertheless concerned with establishing common consent about what is going on in the (apparently!) objective world: what can be observed, repeated, measured, proven (as much as that is possible), and so on. When we study biological systems closely: be it our own, or that of other animals, we can, to our satisfaction (or at least we can anticipate how we *might* find satisfactory answers in the future) account for how those biological systems operate in physical terms (e.g. Chalmers, 1995): for example, how they grow, learn, communicate, evolve, and do all those other things we listed above. If we weren't human ourselves, we might be wholly satisfied with those answers, but in fact we *know* (or at least think we do) that this objective account appears to be missing something out, and furthermore, what it's missing out is, for many people, the most important thing. We encounter the 'mind-body problem', an 'explanatory gap', or the 'hard problem' (e.g. Blackmore, 2003; Chalmers, 1995; Levine, 1983): i.e. the feelings and subjective awareness each of us has about the world around us, and of our own thoughts, motivations and emotions, appears to be the most real thing, yet such experiences *seem* to be qualitatively different from the physical world: i.e. they seem to be different in kind from the only world which scientific proof and measurement pertains to. In fact, we know we're in trouble as soon as we try and explain what it's like to have those experiences: we typically talk in circles, trying to point to such phenomena using the inadequate tools of our spoken language (e.g. Block, 1995), and settling for

---

[1] 'Conscious(ness)', in this sense we'll be using the term in this Introduction, refers to subjective experience, i.e. that there is 'something it is like' to be conscious (after Nagel, 1974). Of course, the word can be used in slightly different contexts (e.g. Blackmore, 2003), for example referring to a state of 'wakefulness' (e.g. as in "the anaesthetic slowly wore off, and the dog regained consciousness"), or 'knowingness' (e.g. as in "I was conscious of the danger I was in, so kept my distance from the dog, and stood near the door!"). There's nothing obviously peculiar and special about these latter contexts: we could perhaps refer to a computer being conscious (as in 'awake'), after we've turned it on and booted it up, and for it to be conscious (as in 'having knowledge') of an attached printer's status (e.g. that it is out of paper); we're making a very different inference if we were to suggest it was conscious in a 'subjective experience' sense, though – i.e. that there is 'something it is like' to be the computer (note, I'm not taking a position on this: there might be!)

(perhaps honest, but unhelpful) terms such as 'ineffable' (e.g. Blackmore, 2003). Good science prides itself on defining, unambiguously, and in fine detail, the subject(s) of its enquiry, none more so than the psychological and behavioural sciences; clearly, if we stumble so badly on that first block, we're off to a peculiarly bad start.

Animal welfare scientists & their problems

For much of science, this doesn't matter a great deal (although, of course, such experiences matter *personally* to the scientists): one can ask questions - and find good answers to them - in academic fields as diverse as biochemistry, cognitive science, computing, physiology, and so on, without encountering such 'explanatory gaps' or 'hard problems'. In animal welfare science, however, the 'problem' matters a great deal; in fact 'animal welfare science', as a discipline would likely not exist at all if subjective experience didn't 'matter' (e.g. Dawkins, 1990; Mendl & Paul, 2004).

To put it in relatively crude terms, whilst we wouldn't think twice about taking an axe and chopping off the legs of an old wooden table to make a bonfire, all else being equal, most of us certainly would think twice before doing the same to a live (non-human) animal, or indeed to each other. So, what's the difference? Well, I personally know that if I were undergoing the 'firewood' treatment (i.e. being chopped up!), I'd find the experience horrible: I'd feel intense pain and despair, and indeed there are few things worse I could anticipate happening! Whilst, as our example above implied, *arguably* I cannot know for sure whether you (or any other human) is actually conscious (insofar as I may not be able to *prove* it), I (perhaps implicitly) note aspects of your anatomy and behaviour (maybe cross-referencing them with my own) which suggest that you, like me, are indeed similarly conscious, and would similarly suffer; as such, I would decline giving you the 'firewood' treatment.[2] By seeing what other objects satisfy my 'it seems conscious' checklist,

[2] Of course, knowing an object is capable of suffering might give some people all the more reason to consign it to the bonfire: perhaps if one wanted to punish it for political transgressions, for example (it's no accident that Guy Fawkes, rather than a wooden table, is the focus of activities on the 5th November!)

I might well extend this net of compassion to many (maybe all) non-human animals, but not to the table, nor to the clockwork toy (although of course others, albeit a minority, might). However, how widely that net is cast is clearly based on a hunch on my part, perhaps combined with an appreciation of the ethical costs of being wrong[3], and with a variety of other factors that inform that hunch (e.g. cultural milieu, etc.) Others, famously, would disagree with my judgement: Descartes, for example, argued all non-humans were non-conscious *automata* (e.g. Descartes, 1968), different in *form* from the clockwork toy (with different behaviours, anatomy, etc.), but equal in every other sense: i.e. the same in *kind,* without any subjective element.

So, as this suggests, people are concerned about the 'welfare' of non-human animals in a manner which differs from their concern about non-living objects, because they (or at least some of them) think that animals may be conscious, as we ourselves are (although the content and structure of this 'consciousness' may, of course, differ between species), and capable of experiencing negative emotional states, such as fear, anxiety, depression (or something akin to them), as well as more positive ones, such as pleasure, contentment, elation, and so on. However, since people differ in their hunches, and how far they would cast this net of compassion themselves, it would be very useful if animal welfare scientists could tell us for sure which animals were, in fact, conscious, so we could all agree, and adjust our ethical debate, and husbandry procedures, accordingly.

As our opening example implies, though, as soon as we ask such questions, an 'explanatory gap' appears before us. As Velmans (2000) notes, "viewed from a first-person perspective, consciousness appears to be necessary for most forms of complex or novel processing. But viewed from a third-person perspective, consciousness does not appear to be necessary for any form of processing" (quoted in Blackmore, 2003). It doesn't (necessarily) follow that being able to

---

[3] For instance, if I had to put a bet on it (and knew nothing 'bad' would happen to sparrows as a consequence), I might say that a sparrow, for example, is not 'conscious': i.e. there is 'nothing it is like' to be a sparrow (after Nagel, 1974); however, if the direction of my bet condemned thousands of sparrows to the bonfire, the ethical implications of being wrong might expand the significance, for me, of the low probability I estimated of 'sparrow consciousness', a probablity which would otherwise have informed my bet.

smile, wake up, talk, remember things (including specific events; e.g. Clayton et al., 2007), decide, know things (even knowing what you don't know; e.g. Smith & Washburn, 2005), run away, learn, and so on, means that one has subjective experience; it's not obvious why all these things could not occur "in the dark" (Chalmers, 1995) (as perhaps they might do for a sophisticated clockwork toy). Similarly, it does not necessarily follow that having a particular brain part, which studies of lesion patients (for example) might suggest to us is important for certain aspects of subjective experience, means an animal (human or non-human) is *necessarily* subjectively-aware.[4]

Towards a pragmatic solution

The 'problem', as we've described it, may be insurmountable for animal welfare science: we might never know, for sure, whether another animal is conscious or not. However, perhaps this shouldn't concern us *too* much, since, as we've noted, we may never know for sure whether another human is conscious or not - we can only ask them about the experiences and feelings they are having, and cross-reference this with our own subjective states, and accept that circumstantial evidence as proof that they are conscious, as we know ourselves to be. So maybe we can build up a convincing dossier of circumstantial evidence that a non-human animal is conscious. That evidence will be rather harder to come by, though, as non-human animals do not have the capacity of verbal report, which is probably the most informative reflection of subjective experience we can get (e.g. Paul et al., 2005).

So, how might we go about building that dossier for non-human animals? Well, as our above discussion suggests, we might look for functional, or anatomical, correlates, whilst keeping in mind that if we find neural structures (for example) in animals which are (anatomically, or functionally) similar to those which seem very closely involved in conscious experience in humans, or perhaps find that animals

[4] These arguments (and one's position with respect to them) relate to 'philosophical zombies': i.e. the 'thought experiment' that a 'person' might exist that was able to do all the things you and I do, behave in exactly the same way (and is perhaps even constructed identically, in certain versions of the argument), but had absolutely no subjective experience (e.g. Moody, 1994).

(or maybe only some species) can do a certain thing that we ourselves can only do consciously (perhaps we might even cross-reference such functional correlates with neural correlates; e.g. Cowey & Stoerig, 1995), it doesn't *necessarily* follow that a particular anatomical structure, or a particular function, could not exist, or occur, without any associated subjective experience at all (as it is equally not obvious with regard to our own neural structures, and functioning); that may seem a pessimistic caveat, but the strength (or otherwise) of such circumstantial evidence may mean the final dossier is very convincing.

So, the state of that dossier, which is being constantly built, and revised (e.g. Butler & Cotterill, 2006; Edelman et al., 2005; Seth et al., 2005; Weiskrantz, 2001), has important implications for animal welfare scientists. Clearly, though, we can't just wait until that dossier delivers a convincing verdict either way before we start concerning ourselves with the welfare of animals, particularly those under our care; rather, we give them the 'benefit of the doubt' (as we do, arguably, with each other), and take the following position: *if* we assume animal 'A' is 'conscious', what can we do which is likely to have a positive impact on that animal's subjective experiences, and what are the best objective measures we can use which can tell us, by (objective) proxy, what animal 'A' is likely to be feeling? These two objectives are closely related, since it's often far from obvious what's 'good' for an animal, in terms of its welfare, and it's very useful to have good 'proxy' measures so that we can cross-reference these with our interventions.

Objective (proxy) measures of subjective states

As such, a range of such proxy measures have been developed over the past few years, and their design partly reflects the observation that in humans, at least, subjective emotional experiences co-vary in characteristic ways with other aspects of our physical functioning. So, typically, when we ourselves *feel* anxious (e.g. perhaps we're in the dentist's waiting room, following a period of dental neglect), certain physiological processes might change (e.g. our heart rate may increase, we may sweat more, certain hormones might be released in greater quantities, etc.), and also aspects of our behaviour might change (e.g. we might seek reassuring shelter behind a magazine in the furthest corner of the waiting room, we might

jump when the receptionist calls our name, we might be very slow to approach the dentist's chair, and so on). Proxy measures of animal emotional states measure objective changes which are somewhat analogous to such human-based indices, albeit developed and interpreted with each particular species' biology in mind. So, there are proxy measures based on physiological change (such as hormonal levels, heart rate, volume of visible eye white, etc.) (e.g. Krohn et al., 2003; Menargues et al., 2008; Sandem et al., 2006), and proxy measures based on behavioural change (approach/withdrawal behaviour, response to novel objects, startle responses, vocalisations, play behaviour, thigmotaxis, and so on) (for reviews see, for example, Ohl, 2003; Paul et al., 2005).

Some limitations of existing proxy measures

Many of these existing proxy measures provide a lot of very valuable information about the state of an animal, but some have important limitations. For example, certain proxy measures may be sensitive to the 'intensity' of an emotional state ('arousal': e.g. feeling 'activated', 'excited', etc., vs. 'relaxed', 'calm', and so on), rather than its 'direction' ('valence'; i.e. whether it's negative or positive). Others measure certain physiological changes in animals which, in humans at least, appear not to be consistently correlated with subjective experience. Sometimes taking the measure itself can change an animal's emotional state, for example if we have to restrain an animal, and/or inject it: in such instances, we may be getting a measure of the animal's response to the measure itself. More generally, there's been greater focus on the development of measures of negative emotional states (e.g. anxiety, fear, depression, etc.), rather than more positive states, even though good welfare might involve both the attenuation of the former, and the enhancement of the latter. In addition, certain measures might be sensitive to relatively acute changes in emotional state, rather than longer-standing, chronic 'moods' or 'affective traits'. Finally, some behavioural measures, in particular, suffer from a lack of clarity regarding *a priori* predictions: i.e. before we conduct a test, we can't unambiguously predict which behavioural result would correspond to an assumed (subjective) emotional experience of given valence and intensity (e.g. Burman et al., 2008b; Mason & Mendl, 1993; Paul et al., 2005).

A 'cognitive', or 'information-processing' approach to non-human animal emotion

In an attempt to address some of these limitations, Paul, Mendl and their colleagues, have recently advocated paying closer attention to the substantial (human) psychological literature relating to subjective emotional experience and certain *cognitive processes,* and to develop and refine proxy measures in light of this (Paul et al., 2005). Some of this literature pertains to the characteristic biases which occur, across emotional state, in the way we process information. In our anecdotal illustration above, in which we felt anxious in the dentist's surgery, we might, for example, have a greater tendency to recall negative memories (e.g. we might remember the last time we made an unpleasant visit to the surgery), our attentional processes might select particular threat-related stimuli on which to focus (e.g. our attention may be more greatly drawn to the drill by the side of the dentist's chair), and the interpretations we make of ambiguous stimuli in our environment might be biased (e.g. we might have a greater tendency to interpret the dentist's laugh following the assurance 'this won't hurt' as sadistic pleasure signifying the very opposite); if such biases characteristically correlate with how we feel, then there is the potential to develop objective measures of emotional state which are faithful correlates of subjective experience.

COGNITION & EMOTION

Definitions of cognition, and how to measure it

So, what is cognition? As Barnard (2004) notes, some conceive of it as "sophisticated information-processing" which a mind, or cleverly-designed computer might do, and might not be consciously-experienced, whereas others, perhaps especially those from an ethological tradition, view it as almost synonymous with consciousness (see, for example, Dawkins, 2001 for a discussion). The former conception is dominant in psychology, and is reflected in Sara Shettleworth's book *Cognition, Evolution and Behavior* (1998), in which she defines cognition as referring to "the mechanisms by which animals acquire, process, store, and act on information from the environment...[including]

perception, learning, memory and decision making" (as quoted in Paul et al., 2005). We favour the latter conception of the word, which does not imply any subjective experience when we talk of 'cognition' (indeed much of human cognitive psychology concerns the study of processes which are not conscious).

How might we measure such 'cognitive' processes? Neuroimaging techniques, at least in humans, are beginning to allow researchers to directly measure the physical changes associated with such processes (e.g. Cabeza & Kingstone, 2006); this is perhaps somewhat akin to opening up a computer, and using devices which allow us to see in which direction, and in which pattern, electrons flow around its circuitry, whilst relating this to the computer's functioning. Such techniques are confined to very specialist situations, and so for practical, everyday purposes, inferring the architecture of cognitive processes in animals (both human and non-human) generally involves the observation of behavioural output. By designing appropriate experiments, and analysing behavioural output carefully, we can infer the nature of the mediating cognitive processes. Similarly, in the absence of a user's manual, we might try to work out how a given software programme processes data by performing designed 'experiments', and observing changes in the computer's output.[5]

Some features of emotions and related phenomena

So, now we've talked a little about cognition, let's also discuss emotions, and related processes, in some more detail. Emotions are evolved phenomena which are adaptive features of the functioning of an organism (e.g. Ekman, 1999). That's not to say that they always operate in a manner likely to enhance an individual's fitness, as certain emotional disorders pay testimony to (e.g. Nettle, 2004). They are elicited by stimuli in the external, and internal environment, and they involve, or are manifested in, changes in a range of biological processes which enable the organism to adaptively respond to the eliciting stimuli, and the change in circumstances they herald (e.g. Rolls, 2005). Those biological processes (or

[5] e.g. if we were unsure whether a computer was checking spelling with reference to an American, or British English dictionary, or weren't sure what base it used when calculating logs, we could find out by 'experimentation'.

'components' or 'aspects' of an emotional response) include subjective experience, physiology, cognition, and behaviour (see Paul et al., 2005, for a discussion of a 'multicomponential' view of emotions). It's important to note that people may differ in which of these components they characterise as being the emotional 'bit': some may think it's *solely* a subjective state (this perhaps best characterises the everyday, 'folk psychology' sense of the word), or *solely* a particular physiological profile, and so on (e.g. Paul & Mendl, in prep.). Such debates are, at least in part, a matter of semantics, and somewhat reflect the academic backgrounds of those involved in such discussions. More practically, whether we were to define emotion as solely subjective, or as involving *all* (or some) of these other components, the changes animals (human and non-human) undergo when in a particular emotional state will still be the same. For our purposes, it is simply important, and useful, to note that a change in a range of processes occurs across emotional state, and some of these might be more (e.g. physiology, behaviour, cognition) or less (subjective experience) amenable to direct scientific enquiry. Since we are privy to our own private experiences, we know that one such component in humans concerns subjective experience; we're less sure whether the same might be true of non-humans, but we can objectively measure the other components and make an informed judgement as to how well these correlate with any conscious experiences that animal might be having.

Bringing emotion back into cognitive science

As our earlier analogy implied, the cognitive approach to the understanding of mind and behaviour owes much to its use of the computer as a metaphor. The 'cognitive revolution', which took place in the 1950's and 60's, superseded the previously dominant doctrine of *behaviourism*. In confining experimentation to what could be objectively measured - namely behaviour ('output') and stimuli ('input') - behaviourism helped to develop psychology as an empirical science (prior to then it had been a more philosophical, introspective endeavour). However, it became apparent that there were important limitations of the behaviourist approach which meant it could never provide a comprehensive theory of mind and behaviour, and it was replaced by a doctrine which allowed the internal architecture of biological information-processing systems (some of which could not be understood

with sole reference to an animal's experience – i.e. the stimuli it had previously been exposed to), to be modeled and tested; importantly, such modeling and testing could occur without a loss of empirical rigor, nor an appeal to introspection.

'Consciousness' and subjective experience are problematic issues for science, and the cognitive revolution owes its considerable success, in part, to being able to circumvent such phenomena: as our 'clockwork toy' anecdote illustrated, by confining our approach to modeling how information is processed (in the example of the clockwork toy: the design of the cogs, the speed with which they rotated, the connections they had with various gears, and so on; in the example of biological systems: the architecture of neuronal networks, patterns of inhibition and excitation, and so on) we can happily proceed without recourse to subjective experience.

Since subjective phenomena are an important part of how we tend to conceptualise 'emotions', they received comparatively little attention by cognitive scientists, at least at the start of the 'cognitive revolution' (e.g. Gray, 1999). In fact, such states were sometimes regarded as nuisance variables: biological phenomena of primitive origin, compromising the rationality of 'pure' cognitive processes, such as thought, attention, perception, and so on (e.g. Loewenstein & Lerner, 2003). More recently, this attitude has changed, and experimental psychology has widened the scope of its enquiry to include 'consciousness', and also emotions, mood and affect (including their subjective components). In an interesting echo of the circumstances surrounding the cognitive revolution, psychologists are increasingly appreciating that the neglect of emotion as a variable means the 'classical' cognitive approach to the understanding of mind and behaviour will only be able to tell part of the story. Importantly, there has been a realisation that by bringing emotion back into psychological theories and models, rather than introducing 'noise' which clouds our ability to clearly see the rational processes of cognition, it may actually *enhance* our understanding of those cognitive processes (e.g. Forgas, 2003; Loewenstein et al., 2001) (much as introducing cognition developed our understanding of key (otherwise inexplicable) phenomena uncovered during the behaviourist period of experimentation). Perhaps coupled with an increasing appreciation that cognitive processes aren't

always 'rational' (e.g. such processes may 'make decisions', for instance, in a manner quite different from how a computer, designed to take full account of, and weigh up, all available information prior to making a 'decision', would do), an understanding of the adaptive significance of emotional states means greater attention has been paid to the role of emotions in 'tuning' cognitive processes in a manner which enables the organism to better respond to a changing (external and internal) world: making better decisions, and so on (e.g. Clore & Huntsinger, 2007; Damasio, 1994; Finucane et al., 2000). More generally, the historical delineation[6] between emotion and cognition is being revised in favour of a view emphasising the integration of the two (e.g. Duncan & Barrett, 2007; Eder et al., 2007; Ochsner & Phelps, 2007; Pessoa, 2008; Phelps, 2006).

Conceptually, the various components of an emotional response could be regarded as different 'bits', or (as a slight variation) as different ways of looking at the same thing; e.g. one could take a 'physiological approach' to emotion (and measure the change in neuronal activity, endocrinological functioning, and so on), one could take a 'cognitive approach' to emotion, and examine changes in the operations, and functioning, of parts of the 'software' running on the neuronal hardware, etc. In part, the merits of a particular 'approach' depend on the types of questions one's asking, and also the scientific progress achieved through adopting a particular paradigm (e.g. Dalgleish, 2003); historically, the 'cognitive approach' has been very successful in both posing scientific questions, and in finding answers to them (e.g. Byrne & Bates, 2006). So a 'cognitive approach'[7] to understanding emotion may serve us well, but of course it won't tell the 'whole' story of 'emotion', nor will it be the best approach in all circumstances.

---

[6] Perhaps originating in classical Greek philosophy: Scherer (2003), for example, notes that "in arguing for a tripartitite model of the soul, Plato created the concepts of 'cognition', 'emotion', and 'conation' (motivation), and placed them in partial opposition to each other." As mentioned in the main text, some argue the historical boundaries between emotion and cognition should now be revised (with some suggesting that 'affect' is a form of 'cognition', for example: e.g. Duncan & Barrett, 2007), plus other aspects of Plato's model are being questioned too, with some proposing that classical 'motivational' states, such as hunger, thirst and pain, may be better conceptualised as emotions (e.g. Tsuchiya & Adolphs, 2007).

[7] Incidentally, some people instead refer to an 'information-processing' approach to emotion (e.g. Dalgleish, 2003); this perhaps allows us to circumvent any debates concerning where (classically) cognitive processes end, and (classically) affective processes start (each of which may involve the processing of information); such debates are historically a prominent feature of psychological theorising (for a review, see e.g., Panksepp, 2003).

Emotions, moods, and affect – some notes on terminology

As Paul et al (2005) note, the words 'affect' and 'emotion' are often employed synonymously (e.g. Pessoa, 2008) as catch-all terms for phenomena such as transient responses to particular emotive events, longer-lasting moods, and so on. We take Paul et al's lead, and use the terms in this interchangeable manner too, but it's important to note that 'emotion' is occasionally used in a more specific sense, to refer to object-related and relatively transient affective states (e.g. an 'emotion' (perhaps anger) elicited in response to an unfamiliar conspecific intruding on one's territory). These are distinguished from 'moods', which are longer-lasting, more diffuse affective states; moods may be precipitated by particular events too, but can then persist for some time in a manner which is not obviously 'object-related', such as a general feeling of depression, anxiety, happiness, etc. (e.g. Clore & Huntsinger, 2007; Paul et al., 2005).

Some researchers stress the categorical differences between emotions, whereas others tend to adopt a more dimensional view (for brief discussions, see e.g. Panayiotou, 2008; Paul & Mendl, in prep.). With regard to the latter, as Scherer (2003) notes, it's a surprisingly pervasive finding that a variety of human affective states can be differentiated by their loci on a number of orthogonal dimensions; these 'axes' variously relate to emotional *valence* (i.e. positive or negative), the direction of *engagement* (i.e. approach or withdrawal), and also *arousal* (i.e. intensity) (e.g. Clark & Watson, 1991; Russell, 1980; Watson et al., 1988). 'Moods' are sometimes conceptualised as affective states which simply vary along these dimensional axes, whilst 'object-specific emotions' might be additionally characterised by certain object-related appraisals (e.g. Paul & Mendl, in prep.). In addition, the term 'affect' is occasionally used to refer specifically to 'valence' (e.g. Paul et al., 2005).

Finally, both 'moods' and 'object-related emotions' tend to be referred to as affective *states*, whereas affective *traits* are a "tendency to react in a particular affective way to a variety of events across time and situations" (e.g. Loewenstein & Lerner, 2003): i.e. something akin to the affective component of a personality.

THE RELATIONSHIP BETWEEN COGNITION AND EMOTION IN HUMANS

The changes in cognitive (or information) processing which occur across emotional state in humans are wide-ranging and complex. Some of these relate to changes in *capacity*, and some relate to changes in *selectivity* (e.g. Dalgleish, 2003).

Affect-related changes in cognitive capacity

An example of affect-related changes in *capacity* might be certain anxious states, in which attentional resources, and a portion of finite-capacity working memory, are more greatly occupied with the monitoring of potential threats in the environment; as a result this may mean that the animal (human or non-human) in question performs more poorly in tasks which are not threat-related (since that animal may focus fewer cognitive resources on such tasks compared to when it is in a less 'anxious' state) (e.g. Eysenck et al., 2007). In another example, when a person is depressed, the capacity and efficacy of a whole range of cognitive processes may be compromised, including memory, learning, decision-making, executive processes, and so on (although, of course, the extent of such impairments depend on the intensity and longevity of the depressive episode; e.g. Gualtieri et al., 2006).

Affect-related changes in cognitive selectivity

Affect-related changes in *selectivity* refer to a biased tendency to process particular types of information in a cognitive system of limited resources (in which *not everything* can be processed); as one of the above examples illustrates, such changes can also affect information-processing capacity. Research on affect-related biases in human information-processing has generally focused on the following areas:

- Memory
- Attention
- Judgement / Interpretation

Soon, we'll discuss these in a little more detail; before we do, though, it's worth noting a few points which will inform that discussion.

Firstly, with notable exceptions, research on the links between cognition and emotion in humans has often focused on negative affective states: specifically anxiety and depression. This is perhaps not surprising, since such states (at least when occurring at a clinical level) are of most medical concern (in a similar way that negative affect is of most concern to animal welfare scientists, and has also received similarly biased attention).

Secondly, a bias in memory (mnemonic) processes, for example, may reflect changes in memory-specific systems, and/or changes in attentional systems, and/or changes in judgement/interpretational systems, and *vice versa*; it's possible that the contribution of these various mechanisms might be distinguished through careful experimental design, but as a general point, saying a human (or non-human) has a specific 'bias in memory' (for example) doesn't necessarily mean that bias solely reflects changes to cognitive systems we generally regard as primarily memory-related.

Thirdly, some researchers are interested in whether such 'cognitive biases' might be found in certain individuals before they develop an emotional disorder which requires clinical intervention, or whether they instead occur *during* that emotional disorder; i.e. whether the biased processing of information might make humans vulnerable to developing affective disorders (in which case it might have significance as a vulnerability marker, and might be the focus of preventative intervention), or whether they only occur once such disorders commence (in which case they might perpetuate such disorders, and may similarly be the focus of intervention) (e.g. Bishop, 2007; Dent & Teasdale, 1988; Mineka et al., 2003; Power, 1999).

Finally, affective states can co-occur. For example, it is often the case that people who are depressed are also anxious: i.e. there is a considerable amount of *co-morbidity* (incidentally, the same is not quite true for those who are anxious: i.e. a greater number of such individuals are without depression than *vice versa*;

furthermore, anxiety tends to come first when they are expressed sequentially) (e.g. reviewed in Mineka et al., 1998). As a general observation, this should alert us to the possibility that if we find that a particular cognitive process is functioning differently in a depressed person (or a non-human who we think might be depressed), that might not reflect their 'depressed state' but might reflect the presence of any co-morbid anxiety (for example).

Below, we will review some of the evidence for certain affect-related cognitive biases in humans.

MEMORY BIASES

There is considerable evidence suggesting that humans can better recall previously-learnt material which has an emotional significance congruent with their current affective state (e.g. Mineka et al., 2003).

Depression

In general, research on mood-biased memory has focused on depression, reflecting a concern that such biases precipitate clinically-depressive states, or maintain them (e.g. Gotlib & Krasnoperova, 1998). Such research has found good evidence that mood-congruent memory biases occur in depression (as reviewed in, for example, Dalgleish & Watts, 1990; Mineka et al., 2003), and that this isn't simply a function of such individuals having more negative memories (i.e. it's not solely due to the possibility that more negative things may have happened to a person who is currently depressed; e.g. Clark & Teasdale, 1982). Furthermore, such depression-related biases are found in 'explicit memory tasks' (e.g. free recall, recognition tests, etc; i.e. when participants are explicitly asked to remember things) (e.g. Bradley et al., 1995; Teasdale & Fogarty, 1979) and also in 'implicit memory tasks' (in which memory is not assessed directly) (e.g. Bradley et al., 1995); such tasks may, in part, map onto different underlying mechanisms (perhaps more 'elaborative', or 'strategic' processing in the case of the former, and more 'automatic' in the case of the latter; e.g. Bradley et al., 1995).

Anxiety

In contrast, the overall evidence of memory biases in anxiety is not as strong (e.g. Russo et al., 1999), although it appears *certain* types of anxiety disorder may be associated with such biases, for example biased retrieval of threat-relevant information in panic disorder, along with evidence, in a limited number of studies, of mood-congruent memory biases in post-traumatic stress disorder (PTSD) and obsessive-compulsive disorder (OCD) (e.g. Coles & Heimberg, 2002).

More positive affective states

Although not as extensively studied, evidence suggests that positive affective states are also associated with mood-congruent memory biases (as reviewed in, e.g. Blaney, 1986; Isen, 1999); more generally, Mineka et al (2003) note that non-depressed controls generally recall more positive than negative material.

ATTENTIONAL BIASES

Anxiety

In contrast to memory biases, there is very good evidence that anxious humans have characteristic *attentional biases* (e.g. Bar-Haim et al., 2007). A number of experimental paradigms have been used to test such biases, including the modified (emotional) Stroop task, and the dot probe task.

Stroop tasks

The Stroop task involves presenting participants with words in different colours. The participant is asked to name the colour of the font in which the word is presented, as quickly as possible. However, the speed with which people can do this very much depends on the nature of the word itself (e.g. what it spells); in the original version of the task, Stroop (1935) found that participants took longer to name the ink colour of a word if it spelt a colour other than that of the ink in which it was printed, suggesting the word's semantic content disrupted, or interfered with, their ability to report on other aspects of the stimuli. Nowadays, there are several

versions of this task in popular use, many of which employ words which spell a variety of carefully chosen non-colour names; by varying the relevance of those words to the study population, such 'modified Stroop tasks' can be used to explore a range of attention-related biases; as such, researchers gain insight into which types of stimuli are more likely to recruit selective attention (and interfere with task performance; e.g. C. MacLeod, 1999).

For example, Mathews & MacLeod (1985) used a modified Stroop task to investigate attentional biases in anxiety; they found that clinically-anxious participants took significantly longer to name the colour of stimuli which spelt threatening words (e.g. 'injury', 'pathetic', etc.) than those which spelt non-threatening words (e.g. 'hobby', 'confident'), compared to participants who were not anxious. Many other studies have found similar effects, with respect to both anxious state, and high-anxious trait, populations (see Williams et al., 1996, for an extensive review). Results such as these suggest that anxious participants' attention is drawn *towards* the semantic content of the threatening words, and away from other aspects of those stimuli (such as their colour), although some have argued that if attention were *diverted* from threat-related stimuli, response latency may be similarly longer (as discussed in C. MacLeod, 1999, for example).

Dot probe tasks

An alternative paradigm, the 'dot probe task', has been employed to address such issues. A typical version of this task involves two stimuli (usually words) being presented on a screen (e.g. one on the left, the other on the right). These stimuli soon disappear, and a 'probe' (e.g. two dots) appears in the position previously occupied by one of the stimuli. The participant is asked to respond as quickly as possible as soon as they detect the probe's appearance (e.g. Bishop, 2007). Again, by varying the relevance of the initially-presented words to the study population, such paradigms can be used to study attention-related biases. For example, MacLeod et al (1986) found that clinically-anxious participants were quicker to correctly indicate they had seen a probe when it replaced a threatening word than when it replaced a non-threatening word, whereas the *reverse* was true of non-anxious participants. Such results (see also, for example, Bradley et al.,

1999; Mogg et al., 1992) suggest that anxious participants *do* in fact have a greater tendency to direct selective attention *towards* threat (rather than to avoid it).

Other characteristics of anxiety-related attentional biases

Such anxiety-related attentional biases are generally specific to negatively-valenced, threat-related stimuli (e.g. Mogg & Bradley, 2005)[8], and the effect tends to be greater if the semantic content of those stimuli is more relevant to the specific concerns of a particular anxious participant (e.g. Mogg et al., 1989)[9]. Furthermore, conditioned, as well as intrinsically-threatening stimuli, can elicit such attentional biases (e.g. Van Damme et al., 2006).

Attentional biases are found with respect to both anxious states, and traits (although the evidence is a little less strong in the case of the latter; e.g. Bishop, 2007), and in clinical and non-clinical populations (e.g. Bar-Haim et al., 2007). Interestingly, clinical populations exhibit attentional biases when the stimuli are presented both subliminally and supraliminally, whereas non-clinical anxious populations tend to show biases more strongly when the stimuli are presented subliminally (as reviewed in, for example, Bar-Haim et al., 2007). This somewhat suggests that there is a bias related to highly-automated, 'pre-conscious' cognitive processes in both study populations, but that sub-clinical participants are better able to compensate using 'higher-level', cognitive control processes: i.e. when participants are consciously aware of stimuli, they may be able to better meet the requirements of the task (to answer correctly as quickly as possible) by adjusting 'higher-level' cognitive mechanisms in the face of 'lower-level' bias (e.g. Mineka et al., 2003; Pessoa, 2008).[10] In clinical populations, on the other hand, such

[8] Although as Mogg & Bradley (2005) note in their review, attentional biases have been found in certain anxious populations with respect to words not obviously threat-related: e.g. 'sadness'.

[9] For example, Mogg et al (1989) employed the modified Stroop task with a clinically-anxious sample of people: participants who reported being mostly worried about health or physical dangers had greater response latencies to threatening words related to such worries (such as 'disease', 'mutilated', etc.), whereas those participants who reported being mostly worried about social concerns had greater response latencies to threatening words related to those social worries (e.g. 'failure', 'inadequate', etc.)

[10] Bishop (2007) discusses the possibility that "selective attention to threat is determined by the relative signal strength from a pre-attentive threat evaluation mechanism versus that from top-down control mechanisms. Anxiety is held to increase the output from the threat evaluation mechanism, biasing attentional competition in a threat-related direction, even when conscious awareness of the threat-related stimulus is absent". She notes that there is considerable evidence that the

'strategic' mechanisms may *also* be biased, or otherwise compromised, so compensatory higher-level processes are either not available, or actively facilitate biased attention (e.g. Bar-Haim et al., 2007; Bishop, 2007; Mineka et al., 2003).[11] Furthermore, there is evidence that attentional biases (especially pre-conscious) may be a vulnerability marker, correlating with a tendency to later develop a clinical condition (e.g. C. MacLeod, 1999; Mineka et al., 2003).

Depression

In contrast to the strong evidence for attentional biases in anxiety, the evidence for attentional biases in depression is less robust (e.g. Mineka et al., 2003; Mogg & Bradley, 2005; Power, 1999); however, when such biases have been found, they tend to occur in studies in which the stimuli are presented supra-liminally, and when those stimuli have greater personal relevance (e.g. Mineka et al., 2003; Mogg & Bradley, 2005). Interestingly, when co-morbid with anxiety, depressed individuals don't always show an attentional bias towards threat (e.g. Mogg & Bradley, 2005).

More positive affective states

With regard to more positive affective states, a recent study has found evidence of an attentional bias towards rewarding stimuli when in a positive mood. Using a dot-probe task, and a variety of positive mood manipulations (including 'natural' variation), Tamir & Robinson (2007) found that positive mood was associated with greater attentional bias towards positively-valenced words which were associated with potential rewards (and were high-arousing; e.g. 'sexy', 'success'), rather than to generally positive (and low arousing) words which were not obviously reward-related (e.g. 'safe', 'carefree'); furthermore, these biases were found when the stimulus words were only presented for a very short time (e.g. 300ms), as well as

---

amygdala plays an important role in pre-attentive threat-related processing, whilst the pre-frontal cortex (PFC) is implicated in top-down, cognitive control.

[11] With regard to how this might map onto corresponding neural substrates, Bishop (2007) notes that "findings provide evidence for anxiety-related frontal hyporesponsivity, as well as amygdala hyper-responsivity, during the regulation of attention to threat-related stimuli".

longer periods (e.g. 900ms), suggesting such biases are relatively 'automatic'. They interpreted these results via a framework describing a *positive affect*-related tendency to approach reward, and a *negative affect*-related tendency to avoid unpleasant outcomes (e.g. Carver, 2001; Watson et al., 1999).

More generally, some have suggested there is a 'default' tendency, in those who aren't currently anxious, to fast-track the processing of possible sources of threat (e.g. LeDoux, 1996), and otherwise a "propensity to attend to, learn from, and use negative information far more than positive information" (Vaish et al., 2008). Interestingly, a recent meta-analysis found no evidence of threat-related biases in non-anxious people (Bar-Haim et al., 2007), which is a curious finding since such a default tendency would, one might think, be adaptive; however, this *might* reflect the attentional bias tasks generally employed in such studies: when there is a *real* possibility of threat, attentional processes in non-anxious people *may* be adaptively deployed towards it, but not towards the sort of stimuli typically used in such studies (e.g. words), which nevertheless attract the attention of anxious people (perhaps reflecting their lower processing thresholds).

JUDGEMENT / INTERPRETIVE BIASES

If there is a certain level of uncertainty, or ambiguity, regarding the significance of a particular stimulus, or regarding the probability that a particular event will occur, some people, on some occasions, are more likely to interpret that significance to be more negative than others, or to judge the occurrence of positive or negative events to be more likely than others. Importantly, such biases co-vary with people's emotional states and traits in a characteristic manner (e.g. Bishop, 2007; A.K. MacLeod, 1999; C. MacLeod, 1999; Mineka et al., 2003).

Studies with humans have found such biases with regard to a number of different stimuli: for example, with regard to words, or sentences, which might have more than one equally valid meaning; with regard to facial expressions which could be interpreted as signifying a person is in one emotional state or another; and with regard to the nature (i.e. positive or negative) of future events, and the likelihood that they will occur.

Biases relating to the interpretation of such stimuli encountered in the here-and-now are generally referred to as judgements of *ambiguity.* Of course, judgements relating to the nature, or likelihood, of future scenarios are, in a sense, judgements of ambiguity too (since the future is unknown), but this fits a little less well with the commonly-agreed semantic interpretation of the term, and so research relating to interpretive biases of 'ambiguity' generally relates to the former scenario.

Interpretations of ambiguity in anxiety

Let's look at some examples of biases relating to ambiguity. Mathews et al (1989) presented clinically-anxious, recovered-anxious and control (i.e. non-anxious) participants with spoken words, and asked them to write them down. Some of these words were homophones: i.e. could be spelt in more than one way, with each spelling having a different semantic meaning; importantly, some of these meanings were more threatening than others (e.g. die/dye, slay/sleigh, and so on). They found that the clinically anxious group were more likely to write down the words using the more threatening spelling than the control group (with the recovered-anxious group intermediate). Eysenck et al (1991) found similar results using whole sentences; they found that clinically-anxious participants were more likely to judge ambiguous sentences (e.g. 'the doctor examined little Emma's growth'; 'the farmer gave Dave the sack'), as having a negative meaning (e.g. 'the doctor looked at little Emma's cancer'; 'the farmer took away Dave's job'), rather than a more benign meaning (e.g. 'the doctor measured little Emma's growth'; 'the farmer handed Dave the bag'), compared to non-anxious and recovered-anxious participants.

It's possible that the anxious participants in these tasks process (and are aware of) both possible interpretations of the stimuli, but are more likely to select, and respond with, the more negative one when asked; alternatively, it may be that *only* the semantic representation relating to the threatening meaning is processed, or reaches 'awareness', and *that* is why they respond with the threatening

interpretation[12] (e.g. Eysenck et al., 1991; Mogg et al., 2006). The former would perhaps best be termed a *response bias*, whereas the latter might more suitably be termed an *interpretative bias* (e.g. C. MacLeod, 1999); whilst eliciting the presence of the former might be of some interest[13], these experiments were chiefly designed to tap the latter, and it is their presence, or otherwise, which is likely to have wider implications for our understanding of affect and cognition.

A number of studies have used a variety of priming techniques to better rule out a 'response bias' interpretation. For example, Richards & French (1992) presented an ambiguous 'priming' word (e.g. 'arms', 'sentence', etc.) to high trait-anxious, and low trait-anxious participants, on a screen. This was soon replaced by a different 'target' word, to which the participant was asked to indicate, as quickly as possible, whether this was an actual ('real') word, or a non-word. In some trials, the real word related to a threatening meaning of the prime (e.g. 'weapons', and 'prison', respectively), whereas in others it related to a non-threatening meaning of the prime (e.g. 'legs', and 'words', respectively). They found that anxious participants were comparatively quicker to respond correctly when ambiguous primes were followed by threatening targets.

Experiments such as these (see also, e.g. Calvo et al., 1994; Macleod & Cohen, 1993), which incidentally employ a number of control stimuli and trials[14], are more likely to tap what we might call interpretive biases, since they explore the *facilitation*, or otherwise (as operationalised by quicker, or slower, response latencies, respectively), of correctly responding to a target stimulus, which presumably reflects the level of activation of particular semantic representations, as manipulated by the priming stimulus. They suggest that, given an ambiguous stimulus with both non-threatening and threatening connotations, a threatening semantic interpretation is more likely to be activated in anxious individuals. By

[12] Or perhaps the interpretation that reaches 'awareness' first.

[13] E.g. if anxious people were more likely to offer a threat-related outlook (outside a lab setting), this may have certain implications for the nature of their social relationships.

[14] E.g. threat-related target words which do not obviously correspond to the 'threatening' meaning of the preceding ambiguous prime (e.g. 'arms' followed by 'stress'), and the use of ambiguous primes with no obvious threatening meaning at all (Richards & French, 1992).

manipulating temporal aspects of the design, such experiments can examine the time-course necessary to elicit such biases (e.g. Calvo & Castillo, 1998; Richards & French, 1992); generally, when such temporal manipulations have been employed, the threat-specific priming found in anxious participants is not found at very short latencies (e.g. 500ms or below) between a prime and subsequent target (although general semantic facilitation is still found, e.g. Richards & French, 1992). This suggests that at least part of the mechanism underlying the anxiety-related interpretative threat bias involves 'non-automatic', 'strategic' processes, which take longer to unfold (e.g. C. MacLeod, 1999).

Such anxiety-related biases are also found with regard to other ambiguous stimuli, such as facial expressions (e.g. Richards et al., 2002; Yoon & Zinbarg, 2007). For instance, Richards et al (2002) found that high trait-anxious participants were more likely to judge faces with (experimentally-manipulated) 'ambiguous' emotional expressions[15] as fearful than low trait-anxious participants[16].

Interpretations of ambiguity in depression

The above findings relate to anxiety, but are 'interpretative biases' also found in depression? The evidence for this is a little less strong, but it does appear to be elicited in certain circumstances. These circumstances are chiefly those that encourage, or allow, elaborative processing, and also those in which the ambiguity in question is self-referential: i.e. when it relates to the participant. For example, Butler & Mathews (1983) presented clinically-depressed participants and non-depressed controls with ambiguous textual scenarios (e.g. "You wake with a start in the middle of the night thinking you heard a noise, but all is quiet. What do you suppose woke you up?"), and then asked them to rank various possible interpretations in the order they would likely come to mind in such a situation; they

---

[15] Incidentally, they weren't more likely to interpret *any* facial expression as fearful, only those which had been morphed from a number of expressions including a fearful one.

[16] Interestingly, following our previous footnote, Bishop (2007) cites evidence suggesting that amygdala activity increases in anxious people when presented with neutral faces that some interpret as threatening, with the prefrontal cortex activated during (top-down) attempts to interpret such ambiguous stimuli in a less threatening manner; as briefly outlined in a footnote, above, these two structures are also implicated in anxiety-related attentional biases.

found that the depressed participants ranked negative interpretations (e.g. "it could be a burglar") higher than the non-depressed group.

More generally, though, a number of other studies which have used experimental paradigms similar to those which have elicited interpretative biases in anxious people, have failed to find such biases in the case of depression. Investigating a clinical population, Mogg et al (2006) found no negative-priming effect of depression on a text comprehension test, using ambiguous sentences[17] (see Bisson & Sears, 2007, for similar results in a non-clinical population; each of these used third-party scenarios). However, they did find depressive effects on a homophone task (i.e. in which participants wrote down words they heard, which could have a number of equally-valid spellings some of which were negative, such as die/dye).

Such findings somewhat suggest that a 'response bias' might be responsible for these depression-related effects, since the experiments which have found 'interpretative biases' have tended to be those more vulnerable to 'response bias' interpretations; however, these experiments have *also* tended to be those which invite self-referential processing, and it's possible that this may be an important determinant of interpretative biases in depression (e.g. Mogg et al., 2006).

In support of this latter position, a study exploring blink reflexes in relation to ambiguous stimuli *has* found evidence of depression-related interpretative biases, which are apparently not due to differences in response bias. Having first established that the magnitude (i.e. 'strength') of peoples' blink reflex is greater when they are asked to imagine scenarios suggested by ambiguous acoustic stimuli, prior to which a negative disambiguating cue had been presented[18], Lawson et al (2002) conducted a further study with participants differing on an

[17] Related to depression-related themes of loss, failure, inadequacy and rejection, such as "Carol felt emotional throughout the service" (e.g. wedding, or funeral?) and "Mandy thought Steve's attitude towards their relationship had changed" (e.g. for the better, or worse?).

[18] They created acoustic stimuli each of which was the merged product of two spoken words, one with a negative and one with a neutral meaning, differing only in the sound of one phoneme (e.g. 'gloom' and 'bloom'). In the first study (sampled from a college population, not tested for depressive symptoms), they presented a 'disambiguating cue' beforehand: e.g. 'low mood' or 'flowering plant' respectively; they presented no such disambiguating cues in the second study (in which they measured blink response in relation to levels of depression).

index of depression; they found that those higher in depression exhibited blink reflexes of stronger magnitude when asked to imagine a scenario invoked by presentation of an ambiguous acoustic stimulus which had both negative, and neutral, interpretations.

Subjective probability judgements in anxiety & depression

With regard to prospective (i.e. future-orientated) cognitions, humans in a negative affective state (or with negative affective traits) tend to judge negative events as more likely to happen to them, and positive events as less likely to happen to them, than humans in a more positive affective state (see A.K. MacLeod, 1999, for a review). There is some evidence that such prospective cognitions differ across *types* of negative affect (or across relevant affective dimensions): for example, some studies have found that anxiety tends to be associated with the *higher* estimation, or generation, of *negative* future events, and depression with the *lower* estimation, or generation, of *positive* events, compared to non-anxious, non-depressed controls (e.g. MacLeod et al., 1997; Stober, 2000).

In a number of experimental papers and reviews, Andrew MacLeod and his colleagues have related such differences in the valency of prospective cognitions across anxiety and depression to the orthogonal affective dimensions of positive affect/activation (PA) and negative affect/activation (NA), as proposed by Tellegen and colleagues.[19] In this formulation, both anxiety and depression are characterised as being high on the dimension of NA, but only depression is characterised as low on the dimension of PA; high NA, they suggest, is associated with higher expectation of negative events, whereas low PA is associated with lower expectation of positive events. There is some empirical support for this position: for example, MacLeod et al (1996) measured participants' PA, NA,

[19] For example, Miles et al (2004) note "Tellegen and colleagues (Tellegen, 1985; Watson et al., 1988) have accounted for the overlap between depression and anxiety by describing each disorder's relationship to two orthogonal factors labelled Positive Affect (PA) and Negative Affect (NA) (Watson & Clark, 1984). PA is a dimension of pleasurable engagement and reflects the extent to which a person feels active, elated, enthusiastic, excited and strong. NA is a dimension of unpleasurable engagement and is characterized by distress, fear, nervousness and anger. Depression is considered to be a mixed state involving high NA and low PA, and anxiety is a pure state involving only high NA, though an extension of the model posits an additional, unique dimension of anxious arousal for anxiety (Clark & Watson, 1991)." Elsewhere, MacLeod (1999) notes that "it has also been suggested that PA and NA relate to reward-driven and punishment-driven motivational systems, such as Gray's (1982) behavioural approach and behavioural inhibition systems"

depression, anxiety, and their subjective probability estimates that various future events (differing in valency) will befall them; factor analyses found two clear factors: one with loadings from NA, anxiety, negative expectancies and depression, and the other factor with loadings from PA and positive expectancies, with negative loadings from depression.

Such differences in the valency of prospective cognitions across anxiety and depression are not *always* found, though: in a study of secondary schoolchildren, for example, MacLeod and his colleagues found that whilst depression and anxiety were associated with the generation of *more negative* future events, neither was associated with the generation of *fewer positive* future events, compared with controls (although their predictions regarding PA and NA were somewhat supported) (Miles et al., 2004). As a result, they speculated that differences regarding positive future cognitions might only be found in those *severely* depressed, or that perhaps they are more closely associated with a 'hopeless-style' (e.g. Abramson et al., 1989) of depression (to which adolescents may be less vulnerable).

Subjective probability judgements in more positive affective states

Many of these studies, above, have compared depressed or anxious populations with non-depressed and non-anxious controls, but how do more positive affective states compare? Nygren et al (1996) induced a positive affective state in their participants by giving them a bag of sweets; when asked to estimate the probability of winning a gambling task, they found that participants who had received the bag of sweets estimated a higher chance of winning than participants who had not. Many other studies have found similar results with reference to positive affect (as reviewed in, for example, Loewenstein & Lerner, 2003), and the general pattern between affect (both positive and negative) and subjective probability estimates appears to be robust.

More generally, the typical finding is for 'control' groups to give higher subjective probability estimates for positive events than negative ones (A.K. MacLeod, 1999);

i.e. there *tends* to be a 'default' optimistic bias (whether this is borne out by reality or not).

UNIVERSITY OF BRISTOL
MEDICAL LIBRARY

Estimating probability with regard to current & past events

In addition, such biases in subjective probability assessments are not confined to future-orientated scenarios[20]. For example, in a famous series of experiments, Johnson & Tversky (1983) gave participants different newspaper reports to read, and then asked them to estimate the frequency with which various fatal accidents occurred in the U.S. Participants who had read stories involving death (either resulting from leukemia, homicide or fire) estimated higher frequencies than participants who had read 'filler' stories not involving death. Contrary to the authors' predictions, this effect was *not* specific to the type of death featured in the news report (e.g. those who read the 'homicide story' were not more likely to estimate higher frequencies for homicide, or other violent endings, compared to other possible causes), but was global, and was related to the change in affective state (as measured by the investigators) induced by the stories.[21]

Some possible mechanisms

How might people, including those who are depressed and/or anxious, make such judgements? MacLeod (1999) discusses a few possible mechanisms; some of these he relates to a variety of judgement heuristics proposed by Tversky & Kahneman (e.g. 1982; 1973), who suggested judgements may be based on retrieval of relevant memories, and/or the simulation of scenarios in which one could imagine such a future event coming to pass. The latter (*simulation heuristic*) might be more likely to occur when no relevant memories are retrievable, with both

---

[20] Incidentally, such biases may be retrospective too, although the mechanisms underlying such 'covariation' biases may be different from future-orientated ones. For example, Tomarken et al (1989) exposed high snake-fearful and low snake-fearful participants to fear-relevant (snakes) and non-fear-relevant (e.g. mushrooms, flowers) pictures, whilst pairing them with aversive or non-aversive stimuli (shocks, tones, or nothing). They found that high snake-fearful participants later over-estimated the frequency with which the snake pictures were paired with the aversive shocks (in fact, the pairing had been random across the various pictures), and this appeared to be related to the aversive nature of the shocks rather than any greater salience they had.

[21] In subsequent experiments, they found that this bias was also found in relation to self-referential, future-orientated probabilities (i.e. how likely do you think X will be the cause of your own death?), and also to mood as induced by stories not relating to accidents at all (e.g. a 'depressing' story of marital break-up and vocational duress; a 'happier' story of luck and success).

(as *availability heuristics*) determined by the availability, or ease, with which relevant memories, or future scenarios, can be recalled, or constructed, respectively. There is good evidence (e.g. A.K. MacLeod, 1999) that both past experience, and future scenario construction, *generally* play an important role in the judgements we make.

With specific regard to affect, though, evidence suggests the difference in prospective cognitions between those who are anxious and/or depressed, and controls, is mirrored by similar differences (across affect) in how well people build such future events into causal narratives (e.g. Byrne & MacLeod, 1997; Kagan et al., 2004).

In addition, there is a correlation between the ease, and pattern, of memory recollection, and judgements pertaining to future events. For example, in one of the studies we cited earlier, as well as asking participants to generate as many future experiences as possible (both positive and negative), MacLeod et al (1997) also asked participants to write down as many past experiences as possible (again, both positive and negative); the pattern of results, across affect, were very similar to the data relating to prospective cognitions (see also Miles et al., 2004, for a similar pattern of results). *If* people do use past experiences to inform the judgements they make about the future, might they use a specific event memory (SEM), or do they refer to a general impression memory (GIM) of past occurrence (A.K. MacLeod, 1999), or both (or neither)? Cropley et al (2000) found that GIM, and not SEM, was associated with subjective probability assessments in 'normal mood' participants[22], but neither was associated with such assessments in depressed individuals; this, they suggest, might reflect a different mechanism which is used, by default, in depressed individuals when making such future-orientated judgements, a mechanism characterised by efficiency (automacity) through prior rehearsal.

---

[22] Although they note that recollection of specific events could well play a role in certain prospective cognitions: e.g. if that event is very recent, or particularly salient.

Somewhat in keeping with this conclusion, Andersen and her colleagues (Andersen & Limpert, 2001; Andersen et al., 1992) have found evidence to suggest that depressives appear to make relatively more *automatic* cognitions (as well as those judgements being relatively more pessimistic and/or less optimistic), compared to those less (or not) depressed, when judging the likelihood of future events. They found that distractor tasks involving attentional load increased the response latencies of depressives *less* than controls when making such judgements; in their later paper (Andersen & Limpert, 2001) they suggested that this reflected greater automaticy of such judgements in depressives, acquired through greater 'practice' or 'rehearsal' resulting from negative rumination regarding future events (for studies examining the relationship between rumination and prospective cognitions, see e.g. Lavender & Watkins, 2004; Lyubomirsky & Nolenhoeksema, 1995).

More generally in this brief review, we are building up a pattern of evidence which supports the functional role of different affective states. For example, anxiety *tends* to be more threat-related and future-orientated, and to involve a relatively larger amount of 'automatic' information-processing; depression, in contrast, tends to more greatly concern loss or failure, is more past-orientated (and/or 'hopeless' about the future), and involves a relatively larger amount of 'elaborative' information-processing. This distinction is not always empirically clear, which may in part reflect levels of co-morbidity, or the over-simplicity of other aspects of the distinction we're making, but such *general* patterns hold. This, in turn, may reflect their respective functional status: anxiety is generally elicited in situations where potential threat is greater, and it may pay to quickly process cues of potential threat quickly, and to pre-empt dangerous future scenarios. Depression, on the other hand, may invite a level of past-orientated reflection on losses and failed ventures, and the re-prioritisation of one's goals in light of that (e.g. Eysenck et al., 2006; Mineka et al., 2003; Power, 1999).

AFFECT AND DECISION-MAKING

Now that we've reviewed some affect-related biases relating to memory, attentional and judgement / interpretative processes, let's widen the picture slightly and

consider relatively more 'downstream' output (Tamir & Robinson, 2007), such as decision-making behaviour.

Some have argued that, in humans at least, current emotions (i.e. those experienced at the time of decision-making), and those emotions anticipated to occur as a result of the various outcomes contingent on one's decision, both play important roles in decision-making. These two affective factors have been variously named (respectively): *anticipatory* and *anticipated* (Loewenstein et al., 2001), *experienced* and *anticipated* (Mellers et al., 1999), and *immediate* and *expected* (Loewenstein & Lerner, 2003). For the purposes of the following section we will adopt the latter terminologies. Figure 1.1 reproduces a schematic diagram of Loewenstein & Lerner's (2003) proposed model, outlining causal influences between various factors implicated in the decision-making process; their conceptualisation won't suit everyone's notion of how affect interacts with decision-making, but it nonetheless provides a useful framework which will allow us to structure our general discussion of those interactions.

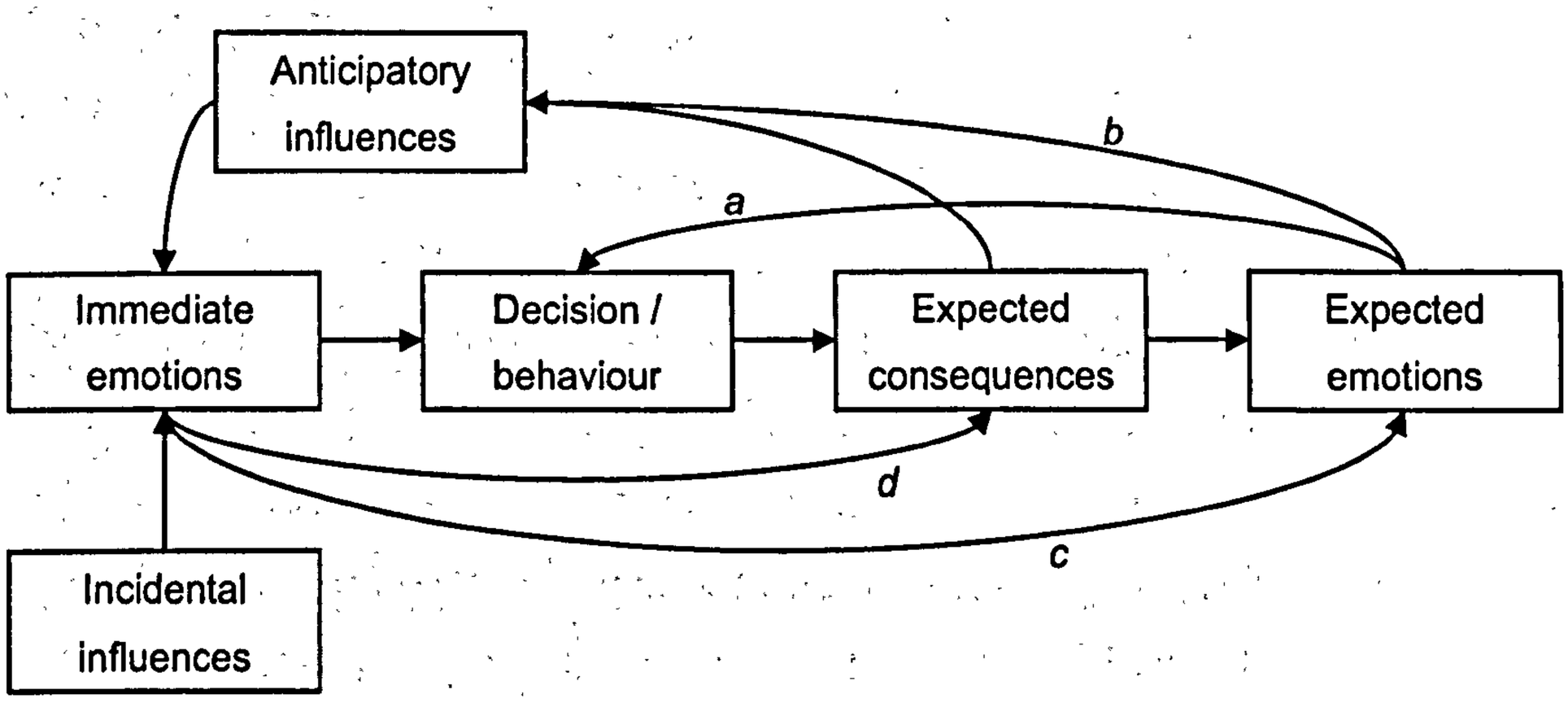

Figure 1.1 Causal influences of immediate and expected emotions (reproduced from Loewenstein & Lerner, 2003).

Expected emotions

The contribution of *expected* (or *anticipated*) *emotions* in human decision-making in part reflects our ability to imagine how we will feel in various future scenarios; we then use this future-orientated simulation, so the theory goes, to try and maximise positive future affect by making appropriate choices now (e.g. Coricelli et al., 2007; Loewenstein & Lerner, 2003; Mellers et al., 1999); this is represented by line *a* in Figure 1.1. The ability for *expected emotions* to influence decision-making in this manner would likely be confined to species able to simulate future events and perhaps 'pre-experience' them, a capacity which *may* be human-specific, or at least specific to species with particularly complex neurological and cognitive architecture (e.g. Atance & O'Neill, 2001). However, it's possible that future outcomes may influence current decision-making in less explicit ways; for example, as a speculative observation, an animal may feel a certain 'contentment' or 'rightness' when building a nest, even though the direct consequences of building a nest may not realised for many weeks to come.[23] Rather than being explicitly acknowledged by the animal, such decision-related outcomes may instead influence its *immediate emotions* through more hard-wired means.

Immediate emotions

With regard to *immediate emotions*, what might determine a (human or non-human) animal's emotional state at 'decision-time'? Loewenstein & Lerner (2003) distinguish *anticipatory influences* from *incidental influences*; the former are anticipatory responses to possible decision outcomes (e.g. positive affect resulting from positively-valenced potential outcomes, and *vice versa*), whereas the latter are extraneous: i.e. not related to the decision at hand. Animals do 'anticipate' outcomes (as illustrated by Pavlov, and many times since), and not in a manner which is solely procedural: adjusting their behaviour in response to changes in the value of those outcomes, for example (e.g. Dickinson & Balleine, 1995). Does

---

[23] As William James remarked, "to the broody hen the notion would probably seem monstrous that there should be a creature in the world to whom a nestful of eggs was not the utterly fascinating and precious and never-to-be-too-much-sat-upon object which it is to her" (quoted in Pinker, 1994).

such an ability reflect the ability to 'pre-experience', or 'imagine', as we discussed above? Possibly, but for many non-human animals, perhaps such a capacity is confined to possible outcomes which will occur in the very near future – e.g. within a few seconds (e.g. Mendl & Paul, 2008) (see lines *a* and *b*, in the diagram).

As well as distinguishing their antecedents, Loewenstein & Lerner (2003) also delineate *indirect* and *direct* effects of *immediate emotions* on decision-making. With regard to the former, as well as influencing our ability to simulate *expected emotions*, and perhaps influencing the subjective value of those outcomes (line *c*, in Figure 1.1), current affect co-varies with changes in information-processing: i.e. its capacity, and selectivity, as we have discussed. This in turn can influence decisions in a variety of ways: via changes in attentional systems, memory status, and judgements of probability (line *d*), amongst others.

With regard to the *direct* effects of immediate emotion, this may be partly determined by *intensity*. At low, or moderate levels of intensity, emotions may take on an 'advisory role' (e.g. the 'affect as information' hypothesis proprosed by Schwarz & Clore, 1983)[24], with decision-makers, in effect, asking themselves 'how do I feel about it?'. In humans, at least, the likelihood that such a ('mood advisory') heuristic will be employed will partly depend on the nature of the decision itself (for instance, whether the decision relates to a familiar or unfamiliar situation, e.g. Forgas, 2003).

At high intensities, on the other hand, emotions may 'overwhelm' classically 'rational' cognition (e.g. Rolls, 1999): phobias might be a good example; for instance, an arachnophobic might *know* (rationally) that spiders (in Britain, at least) aren't harmful, yet may very well be overcome with an aversive emotional reaction when confronted with one. In addition, perhaps more so at higher intensities, emotions bring with them 'action tendencies' (e.g. Frijda, 1994): one might be more

[24] In a famous study, Schwarz & Clore (1983) asked people, on sunny, or rainy days, how satisfied they were with their lives; their judgements were correlated with the putative mood-inducing qualities of the weather (i.e. more satisfied when sunnier), however when their attention was first directed to the weather, the 'rainy day' effect was abolished (although the 'sunny day' effect was not). This suggests that (mis)attributing the source of one's current affective state can be an important determinant of whether such an 'affect-as-information' heuristic is used, and any 'carryover' moods resulting from other incidents *could* play an important role in situations about which the mood has little obvious relevance (e.g. Loewenstein & Lerner, 2003).

likely to hit when angry, run when scared, and so on. Emphasising the adaptive significance of such affect-related tendencies, Loewenstein & Lerner (2003) note that "in this view, emotions save cognitive processing by triggering time-tested responses to universal experiences (such as loss, injustice and threat)." As a more general point, emotions may provide the 'motivation' to take certain adaptive courses of action: courses of action which may not otherwise be taken.

Decision-making under risk

Some have argued that a more explicit consideration of affective factors will likely improve (and has already improved) many classical models of decision-making, including those pertaining to decision-making under risk (e.g. Loewenstein et al., 2001; Mellers et al., 1999; Slovic et al., 2004). For example, expected utility theories view decision-making as a function of the utility (or desirability) of various outcomes, and the probability of attaining those outcomes (for example, von Neumann & Morgenstern (1947), cited by Eysenck & Keane (2005)). From such functions, a number of axioms have been derived of how people should behave if the decisions they make proceed on an optimal basis (as modeled by such theories; e.g. Hastie, 2001); however, important violations of such axioms may, in part, be accounted for by more explicit reference to emotional influences (Loewenstein & Lerner, 2003).

We've encountered many of these in our discussion so far: perhaps most obviously the observation that subjective probability assessments are influenced by current emotional state. Interestingly, though, biased subjective probability assessments don't always influence decision-making behaviour in an obvious manner. For example, whilst Nygren et al (1996) found that (experimentally-manipulated) positive mood increased participants' subjective probability assessments of positive outcomes upwards (i.e. in an optimistic direction), compared to no-mood-manipulation controls, they were not *necessarily* more likely to gamble in a situation which could incur real monetary loss; in fact, they found that when in a positive mood, participants were *more* likely to gamble when potential losses were small (but the probability was relatively large), but *less* likely to gamble when this situation was reversed (i.e. when potential losses were relatively larger, but the

probability of that occurring smaller). Here, and elsewhere (Isen, 1999; Isen et al., 1988), they have accounted for such findings by suggesting that positive mood shifts greater decision-making weight onto the utility (i.e. outcome) of a choice, and away from the probability of that outcome coming to pass. They propose that sensitivity to potential losses is greater when one is in a more positive mood, and that such 'cautious optimism' is adaptive in one with more to lose (i.e. insofar as having more to lose is reflected in one's positive mood).[25] With regard to more negative affective states, the evidence for risk-taking, or otherwise, in depression is fairly equivocal (e.g. Hockey et al., 2000; Mitte, 2007; Yuen & Lee, 2003), but anxiety, as perhaps one might expect, is commonly associated with risk-aversion (e.g. Maner et al., 2007; Maner & Schmidt, 2006; Mitte, 2007).

A more explicit consideration of the role of affect may also help us understand certain impulsive behaviours (i.e. when short-term gain may benefit at the expense of long-term interests). In general, animals prefer shorter, rather than longer, delays to rewards, with an outcome delay of the same magnitude having a greater bearing on decision-making if it occurs sooner (e.g. reward arriving in 5 or 10 seconds) rather than later (e.g. reward arriving in 35 or 40 seconds) (e.g. Ainslie, 1975). Introducing certain affective factors into such models may help better explain various impulsive behaviours: for example, mood can influence estimates of temporal duration, and thus the perceived extent of delays (e.g. Wittmann & Paulus, 2008). In addition, when an outcome is likely to happen very soon, and/or when one has sensory contact with it, for example, 'impulsive' behaviour is more probable (e.g. Hoch & Loewenstein, 1991). Furthermore, certain emotional states (e.g. hunger, sexual arousal), and certain emotional intensities (e.g. very fearful, or very angry), are associated with a higher likelihood of 'impulsive' behaviour (e.g. Loewenstein, 1996). So, for instance, if one is very angry, and there is a face to punch, one might (perhaps inadvisably) punch that face (especially if it belongs to the person who made you angry): here, short-term benefits (of retribution, for example) may receive disproportionate weight in preference to long-term costs

[25] More recently, Kliger & Levy (2003) have found a similar correlation between putative mood (as indexed by weather conditions, e.g. cloudy, sunny, etc.) and risk-related behaviour in the capitals market (a real world situation where potential losses are, presumably, very large), with better moods associated with greater risk aversion.

(e.g. of police proceedings): this differential weighting was part determined by (a) your emotional state and (b) the nature and (both temporal and physical) proximity of the outcome; if these had been different, then the consequences may have been weighted differently.

A NOTE ON BIAS AND DISTORTION

Given that different people, at different times, differ in how they interpret and judge the world around them, and that this co-varies with affect, are there any affective states and traits which are 'most right': i.e. which are associated with a judgement and interpretation of the world which is (closest to) an accurate representation of the objective facts? For many of the experimental paradigms we've outlined above, there is no 'right' or 'wrong' answer: e.g. an ambiguous sentence really *could* mean either of two (or more) different things, at least without further contextual cues. Power (1999) calls this observation (i.e. that there are many valid ways of looking at the same thing) the "Rashomon effect" after Akira Kurosawa's film in which various protagonists offer different versions of the same event.

In other circumstances, however, it might be possible to say whether a given 'bias' is 'realistic' or not. For example, if we're more likely to interpret the emotional significance of facial expressions one way or another, or more likely to be optimistic, or pessimistic, regarding our chances of winning the lottery (for example), these biases *can* be judged against the facts of the matter at hand: e.g. that particular person really *was* experiencing one emotion, or another, when we saw their face (putting aside the possibility they might be e.g. angry and surprised at the same time!), and the statistical chances of winning the lottery on any given week are actually known (at least by someone). Some have suggested there are 'optimistic biases' or 'positive illusions' in 'normal, healthy' individuals (e.g. a default tendency to over-estimate the chances of positive outcomes), and a 'depressive realism' in those experiencing more negative moods (i.e. a depressive outlook which, whilst comparatively pessimistic, is closer to the 'truth') (for various studies & discussions, see e.g. Ackermann & Derubeis, 1991; Lench & Ditto, 2008; Power, 1999; Taylor & Brown, 1988). However, there are plenty of studies which have found this *not* to be the case (e.g. Ben Mansour et al., 2006; Moore & Fresco,

2007; Shrauger et al., 1998), and this more generally illustrates the point that the relationship between 'realism' and 'bias' rather depends on the conditions in which that relationship is tested. For example, Power (1999) distinguishes 'bias' from 'distortion', noting that "a bias is a proclivity to take one direction over another which, under some conditions, will lead to accuracy or realism, but under other conditions will lead to inaccuracy or distortion. In contrast, distortion is invariably wrong".

CONCLUDING COMMENTS

So, we have discussed, in this Introduction, the need for animal welfare science to develop objective indicators which are as faithful a correlate as possible to any subjective emotional experiences an animal might be having. To this end, a number of such measures have been developed; these provide animal welfare scientists with very valuable information, but have some important limitations (e.g. Mason & Mendl, 1993; Paul et al., 2005). In an attempt to address such limitations, Paul et al (2005) have recently proposed investigating the relationship between cognition and emotion in non-human animals. The majority of the scientific literature pertaining to this relationship has been based on studies of humans, portions of which we have reviewed in this Introduction. Hypotheses can be generated from this literature which can, potentially at least, be tested in non-human animals (Paul et al., 2005).

In their review, Paul et al (2005) note that a dissociation between certain physiological components of affect, and subjective experience, has been reported in a number of studies, including those with alexithymics (e.g. Lane et al., 1997; Stone & Nielson, 2001).[26] If we were to develop proxy measures of subjective affect in animals which were sensitive to changes in some of the cognitive processes we've discussed, are there grounds for supposing they might be a more faithful proxy?

[26] Scherer notes that "the apparent dissociation between objective physiological indicators and subjective reports of affective experience (e.g. Myrtek & Brugner, 1996)...might well be due to the fact that only relatively weak affective experiences, for which one might expect little synchronization of the different emotion components (see Scherer, 2001) have been studied so far"; I.e. to the relative insensitivity of certain measures.

Some mechanistic explanations have suggested that felt emotional experiences may have a direct effect on decisions and judgements (e.g. the affect as information hypothesis (Schwarz & Clore, 1983) we discussed earlier, and also the somatic marker hypothesis (Damasio, 1994), and affect infusion model (Forgas, 1995)). As the initial part of our Introduction suggested, however, even if we were to find that non-human animals make decisions, across putative affect, in a manner superficially similar to the way in which humans characteristically do when employing felt states as an informational resource, that doesn't necessarily mean that non-humans are doing the same (although with the accrual of circumstantial evidence, we might accept this was likely to be the case). In addition, proving that subjective states are causal (as opposed to epiphenomenological) factors in human decision-making is a challenging, perhaps intractable, matter (e.g. Castiello et al., 1991; Libet, 1985; Paulignan et al., 1990; Wegner, 2003).

More generally, though, many (if not all) of the papers we have cited which have explored affect-related changes in such processes have measured affect using self-report instruments (i.e. asking participants, in effect, how they feel; or have used clinical populations whose diagnosis would have been based, in part, on such self-report measures). Thus even when some of the biases we have reviewed have involved 'pre-conscious', 'automatic' processes, those biases nevertheless have correlated with a felt state.

Is this correlation *always* found, though? A study by Winkielman et al (2005) suggests not. Using subtle means (subliminal presentation of stimuli), they induced affective change in participants which they were (apparently) not aware of, but which were nevertheless reflected in subsequent preference behaviour (the amount of a novel drink poured and drank), and also value judgements (how much they were willing to pay for the drink). Concluding, as they did, that 'unconscious affect' was elicited in this experiment rather depends on whether the self-report measure they used was sensitive enough to any felt change, but their study at

least suggests that cognitive changes *may* reflect changes in emotional states about which a person (or non-human) is not aware.[27]

More generally, it's likely there are some aspects of cognition which are heavily involved in emotional processes (and functionally vary across them), and others which are relatively (or perhaps totally) 'insulated' (Dalgleish, 2003) – i.e. which operate in just the same manner no matter what the affective status of the animal is. Somewhat similarly, a considerable amount of cognitive psychology involves investigating (cognitive) processes which *may* not (i.e. they can operate 'on-line' or 'off-line'), or *cannot*, be subjectively experienced. However, to the extent that our subjective experience has an informational aspect, aspects of our cognition may faithfully correlate with our consciousness.[28] More generally, increasing investigation of, and theorising about, the relationship between cognition, emotion and subjective experience will likely chart this territory in a little more detail (e.g. Paul & Mendl, in prep.; Tsuchiya & Adolphs, 2007).

As we have seen, many of the existing experimental paradigms have been based on verbal tasks, and therefore, to study affect-related changes in information-processing in non-humans, there is a need to develop novel non-verbal tests (Paul et al., 2005). To this end, Harding et al (2004) investigated 'cognitive bias' in rodents, employing a novel paradigm. Since then, and during the preparation of this thesis, a number of other studies, informed by a 'cognitive bias' approach to the investigation of affect in non-human animals, have also been published. We will discuss some of these in our general discussion, but will introduce Harding et al's study in some detail in the next chapter, before describing three experiments, over the next three chapters, in which we adapt aspects of their paradigm. Since the first three experimental chapters (2-4) are closely related, we provide quite a lot of introductory information in the first of these (Chapter 2), and later refer back to this when describing our subsequent studies; as such, Chapters 3 & 4 are

[27] See also Adolphs et al (2005) for a study of a lesioned patient who expressed a strong preference between different drinks, without being able to perceive their taste, nor being consciously aware of any affective reaction to them (further discussed in Tsuchiya & Adolphs, 2007).

[28] Chalmers (1996) calls such a cognitive correlate of 'consciousness' 'awareness': remarking that 'consciousness' is how mind feels, whilst 'awareness' (as he defines it) is what mind (causally and functionally) does.

considerably shorter than Chapter 2. In our final experimental chapter (5), we introduce a new paradigm, with a different study species, before concluding, in the final chapter (6), with a general discussion.

As a concluding comment, occasionally in this thesis, we will make predictions specific to one species derived from the literature pertaining to another (most often humans); a number of such predictions were implicit in our above discussion (although we weren't making such inferences at all times), but will be made more explicit in subsequent chapters. These predictions will generally concern the relationship between affect and cognition, and they will not be accompanied with detailed discussion of any analogous anatomical features to support such inter-specific speculation. Whilst more generally, the similarities in emotion-related anatomy, physiology, and behavioural processes between species is very impressive (e.g. Berridge, 2003; Blanchard et al., 2001; Haug & Whalen, 1999), considerations of anatomy and physiology are not the main theme of this thesis, and we don't justify every cross-species prediction we make by relating it to anatomical plausibility. Instead, the plausibility will be derived from *a priori* functional considerations of how human-specific certain relationships between affect, cognition and behaviour, are likely to be (as, indeed, we have already discussed in places). Operationally, all the study species employed in this thesis have discrete cognitive processes, allowing them to remember, attend to things, interpret the significance of stimuli, make decisions, and so on. Furthermore, all have evolved in an environment where potential threat will likely bring about an adaptive response akin to some aspects of human anxiety. Moreover, that environment will differ in important ways to the one in which they are presently captive; likewise the environment in which modern humans evolved differs in important ways to the one in which most of us now find ourselves; that general observation leaves open the possibility that relatively basic, adaptive, processes resulting from a persistent thwarting of motivations, for example, might manifest in more intensely depressive states, or their more frequent occurrence, in an phylogenically unprecedented environment. It is these somewhat substrate-neutral observations, and an approach which stresses the information-processing aspects of basic functional biology, which inform the inter-specific predictions we make, always with the caveat that they, of course, may prove ill-specified.

CHAPTER 2

UNPREDICTABLE HOUSING AND JUDGEMENTS OF AMBIGUITY IN RATS

INTRODUCTION

As briefly mentioned towards the end of the preceding section, Harding et al (2004) recently developed a novel (non-verbal) paradigm designed to measure affect-related biases in certain cognitive processes in a non-human animal. In this chapter we provide a summary of their study, outline certain issues raised by their methodology, and then discuss how these might be addressed by alternative designs, such as the one we subsequently employ.

Harding et al (2004) trained rats to perform two different behavioural responses when presented with two different stimuli, using a 'go/no-go' design in which the reinforcement schedules differed qualitatively from each other. In a 'go/no-go' design, animals are trained to perform a behaviour (to 'go') following the presentation of a particular stimulus, and to *refrain* from performing that behaviour (to 'no-go') when presented with a different stimulus; thus, they demonstrate their ability to discriminate between two different stimuli by either engaging, or not engaging, in a particular behaviour.

In their task, the rats were trained to discriminate between auditory tones of 2kHz and 4kHz frequency (we'll call these the 'reference' tones); the 'go' response was pressing a lever which, when correct (i.e. when elicited in response to the auditory tone the experimenter had assigned to it), was reinforced by the delivery of one pellet of food. When the lever press response was incorrect (i.e. when it followed the 'wrong' tone), it resulted in the presentation of 30 seconds of white noise, and a longer fixed interval until the next trial (with its possibility of food reward; the assignment of the two different tones to the 'go' or 'no-go' responses were counterbalanced across rats, but the same for each rat). So, the delivery of food positively reinforced lever pressing following the presentation of one of the auditory tones, whilst avoidance of the white noise stimulus, or avoidance of the longer

interval before the next trial (or both), reinforced the 'no-go' response of not pressing the lever, following the presentation of the *other* auditory tone (alternatively, or in addition, the rat might have simply learnt to *not* associate that particular stimulus with the receipt of food).[29]

When the rats had learnt the task well, approximately half underwent a treatment designed to induce a negative change in their affective state; this consisted of a series of housing events designed to be mildly stressful (such as a reversal of the light cycle, exposure to an unfamiliar conspecific, and so on), delivered on an unpredictable schedule. For the remaining rats, in the control group, the husbandry regime continued as it had done before (i.e. relatively 'predictable', without these events). This treatment was based on the chronic mild stress (CMS) procedure developed by Willner and his colleagues to produce 'anhedonic'[30] states in rodents. The CMS procedure involves the application of a range of aversive events, delivered on an unpredictable schedule, and as such models some of the antecedents implicated in the development of human depression (i.e. experiencing stress over a chronic period in a manner which is unpredictable, and uncontrollable; e.g. Cabib & PuglisiAllegra, 1996; Willner et al., 1992).

The treatment lasted for 19 days, during which they received five further training sessions, before undergoing daily probe-testing sessions over the last ten days of the treatment. In these probe-tests, as well presentation of the 2kHz and 4kHz reference tones to which they had been previously trained to respond, three different tones of a frequency intermediate to these stimuli were also presented: namely 2.5kHz, 3.0kHz, and 3.5kHz; responses to these probe stimuli were not reinforced.

Since, as we discussed in the last chapter, humans in a more negative affective state are more likely to interpret ambiguous stimuli as having negative significance,

---

[29] Or perhaps *not* learn to associate (i.e. low excitation, rather than high inhibition, of the corresponding associative neural networks).

[30] "The decreased capacity to experience pleasure of any sort" (Fawcett et al., 1983; cited in Willner, 1997); elsewhere described as a "subsensitivity to reward" (D'Haenen & Andrews, 2000). In the CMS model, anhedonia is operationalised by a decreased intake of sucrose (or saccharin) solution (e.g. Willner, 2005).

and are more likely to estimate a higher probability of negative events occuring (and/or estimate a lower probability of positive events occuring) than humans in a more positive affective state, the authors hypothesised that rats in the unpredictable housing treatment would be more likely to refrain from pressing the lever in response to these probe tones: i.e. to respond to these stimuli *as if* judging them to be the tones associated with the more negative (or less positive) outcome of the white noise stimulus. Furthermore, if the rats in this treatment group *did* press the lever following a probe tone, they were predicted to do so more slowly: i.e. to have a longer latency.

Their results provided support for these hypotheses (see Figure 2.1 for a copy of their relevant charts): compared to the controls, the rats in the unpredictable housing treatment were significantly slower to press the lever following the food-associated tone, and the probe tones near it, during the probe test sessions ($p<0.05$). There was also a non-significant trend for the rats in the unpredictable housing treatment group to make fewer lever presses when presented with these particular tones ($p=0.10$). A number of concurrent tests were also conducted with the subjects, including an elevated plus maze test, a 'holeboard' test, a sucrose preference test, a food consumption test (time taken to eat 50 pellets of food), and also post-mortem measures of aspects of their neurophysiology and neuroanatomy implicated in stress-related responses. These tests found no significant differences across treatment, apart from the holeboard test, in which the unpredictably-housed group made significantly more rears ($p=0.038$).

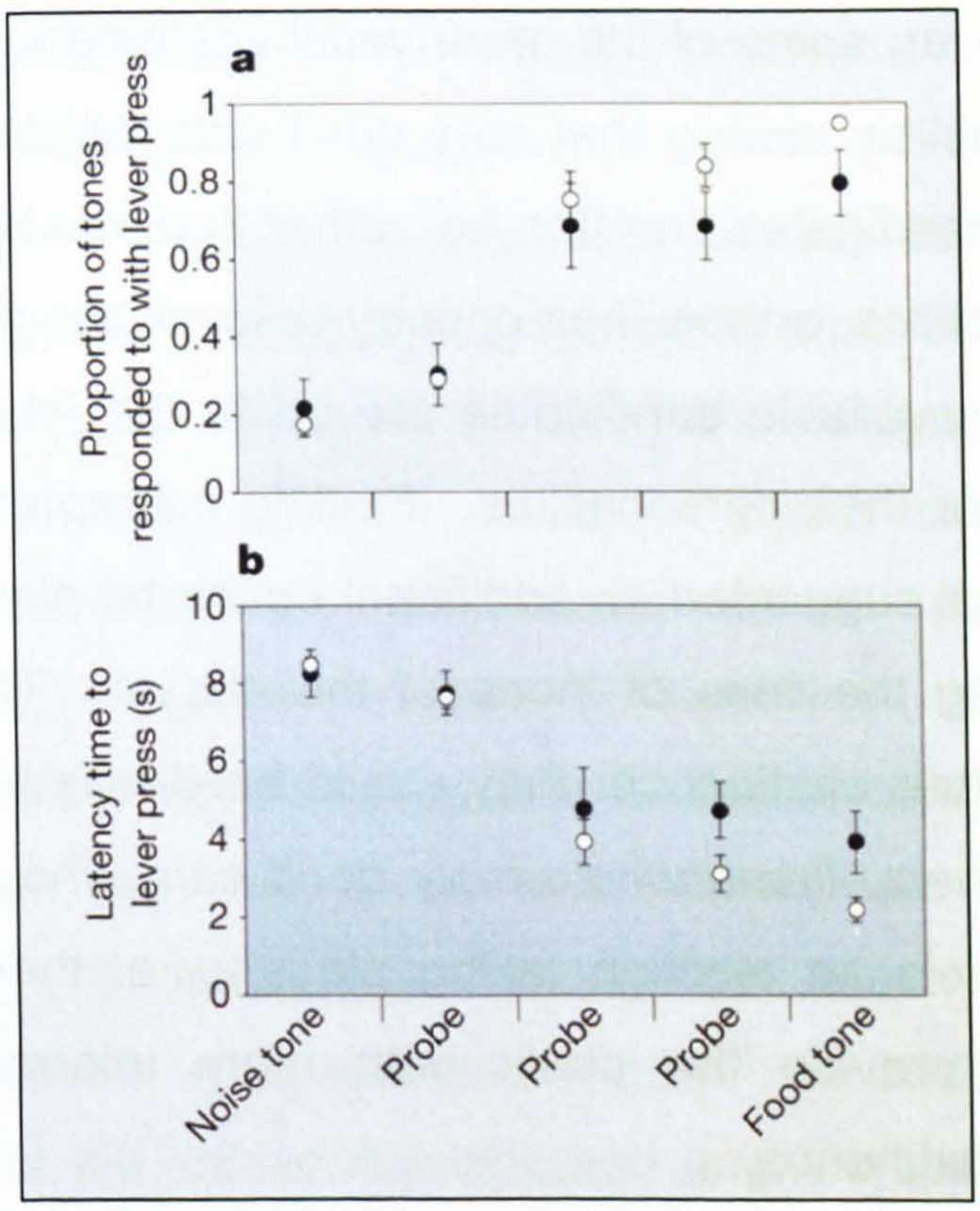

Figure 2.1 A copy of the plot of results from Harding et al (2004), charting the proportion of lever pressing across probe value in **a**, and the latency to record a lever press across probe value in **b**. Open circles = control group; filled (black) circles = unpredictably-housed group (+/- 1SEM).

Clearly, this was an encouraging result, using a highly innovative new paradigm. However, the sample size was relatively low (n=9, in a between-subjects design), not all the predictions were satisfied by the findings (e.g. the difference in response choice was not significant), and the unbalanced 'go/no-go' design, whilst chosen for good reasons, introduces difficulties when attempting to interpret the (biological) significance of the results.

An anecdotal example might help us consider this further: the electoral voting system in the UK involves making one's choice on a ballot paper on which a number of candidates are listed. The level of voter participation has steadily declined over the past few years so that, increasingly, those given the right to vote are not turning up at ballot stations to do so. As one might expect, there has been a considerable amount of speculation as to why this might be the case, with some suggesting would-be voters are becoming increasingly apathetic: perhaps not caring who is elected to office, or at least not caring enough to make the journey to the voting station, or perhaps completely oblivious to the fact that there is an

election at all. However, some of the more vocal abstainers have taken affront at such suggestions, publicly stating that *they* don't vote because they don't favour any of the particular candidates on offer; i.e. rather than not being engaged, or not being interested, in politics, or how their country or locality is governed, they simply don't feel any of the available candidates are up to the job: perhaps they don't favour their policies, or their personalities. Feeling misrepresented by the voting system, they have thus suggested an additional option be placed on the bottom of the ballot paper, along the lines of 'none of the above'. This, they feel, would clarify the nature of their abstinence: they would be distilled from the larger pool containing, among others, those who simply don't care who represents them (or don't even know there's an election taking place!); i.e. they favour a modified design which would provide the public with more information regarding the significance of such abstinence.

Whilst not seamlessly analogous, this rather tangential aside nevertheless gives us more of an intuitive notion as to pitfalls of interpreting 'generically doing nothing' as 'specifically meaning something'. In a 'go/no-go' design, rats making a 'no-go' response may, like an apathetic voter, be simply uninterested in the task (perhaps the possible consequences of engaging with it are less appealing to them, or they are generally less attentive or engaged), or they may be less active (perhaps less willing to make the journey to the lever, and press it; perhaps they generally move around the operant chamber less). In the case of Harding et al's (2004) study, then, there's a danger of erroneously concluding that a rat making a 'no-go' response is categorising a probe stimulus as the 'white-noise-associated' tone (or making a similar interpretation along these lines), when they simply might be less active, or less interested or engaged (in fact, they may even be 'categorising' the probe stimulus as the 'food-associated' tone); i.e. these alternative explanations cloud our ability to unambiguously interpret the findings.

Such confounds are of particular concern in an experiment employing a treatment designed to induce a change in affective state: as well as the characteristic biasing of certain aspects of information-processing, such changes are also, of course, correlated with, or manifested in, characteristic changes in a suite of other biological systems (e.g. Paul et al., 2005). Some of the most commonly-used

criteria for diagnosing depression and anxiety in humans, for example, includes changes in motor activity and appetite (in either direction), interest and concentration (typically lowered, in the case of depression), anhedonia (in depression), and so on (*Diagnostic and Statistical Manual of Mental Disorders IV*, 1994). Likewise, animals exposed to stressors designed to induce states analogous to certain human affective disorders exhibit changes in locomotory activity, sensitivity to putatively pleasurable stimuli ('anhedonia'), feeding behaviour, etc. (e.g. D'Haenen & Andrews, 2000; Willner, 1997). Aware of such confounds, Harding et al (2004) employed a number of concurrent tests designed to gauge such changes, but whilst such supplementary data is undoubtedly useful, such measures may not be sensitive enough to detect any real differences, especially if the treatment is comparatively mild[31]; in addition, the validity and reliability of such tests are sometimes disputed (e.g. Forbes et al., 1996; Reid et al., 1997), and when differences are detected (as in the holeboard test they employed) the (biological) significance of such findings are not always clear (e.g. Paul et al., 2005).

To eliminate such confounds (or at least reduce their likely effect), one would usually employ comprehensive counterbalancing (e.g. Martin & Bateson, 1993). Using the example of the design employed by Harding et al (2004), the behaviours the animals are trained to perform (or not to perform) in response to the conditional stimuli, would normally be counterbalanced: i.e. whilst for half the subjects the contingencies would remain as described above, for the other half they would be reversed: i.e. the subjects would refrain from pressing the lever (i.e. 'no-go') to receive food following the corresponding stimulus, and press the lever (i.e. 'go') to avoid the white noise stimulus following the tone associated with that outcome. In practice though, it would likely prove very difficult to train rats in such a procedure, at least within a realistic timescale. Indeed, Harding (2002) reports a pilot study in which she attempted to train rats in such a task: rather than pressing a lever to stop, or avoid, the presentation of white noise, though, despite Harding's best efforts, the subjects simply 'sat it out'. It may be possible to train rats to make an

[31] The treatment employed by Harding et al was a considerably modified version of a putatively more intense procedure (e.g. Willner et al., 1987).

UNIVERSITY OF BRISTOL LIBRARY

MEDICAL

active response to avoid a particular event if that stimulus was more aversive[32]; however, this would, of course, take the experiment down a road in which the ethical balance of the research programme becomes increasingly unstable. Otherwise, one could consider using conditional responses more closely-aligned to rats' untrained response to aversive events (e.g. Shettleworth, 1998), such as escaping into a recessed hole; however, such qualitative differences in reinforcement still leave the design vulnerable to confounds introduced by any changes in the utility of the unconditional stimuli across treatment: i.e. in how much the experimental subjects value them, or seek to avoid them.

Here, we attempt to address these concerns by adapting aspects of Harding et al's (2004) methodology in an alternative, two-choice design (as indeed recommended by the authors themselves in the conclusions to their paper). The protocol we employ here also involves training rats to press a lever following presentation of a particular auditory tone to receive one pellet of food. However, when presented with a different auditory tone, the subjects are trained to press a *different* lever to receive two pellets of food.

In an earlier pilot study (see Appendix A, p.310), we found that rats preferred to press a lever reinforced with two pellets of food, than to press one reinforced with one pellet of food. This intuitively reasonable finding agrees with a variety of other studies, which have also found similar patterns of preferential responding across reinforcer magnitude in rodents (e.g. Ito, 1985; Reed, 1991); for example, Logan (1965) found that rats preferred to make a choice reinforced with a larger amount of food even when the delay until post-choice receipt of that food was greater than that following the alternative choice reinforced with a smaller food amount; he found that the larger the difference in reinforcer magnitude, the larger the delay

[32] Whilst in pilot testing, Harding (2002) found that the rats did exhibit some freezing and flight behaviour when initially presented with white noise at 70dB, elsewhere duration of white noise at 75dB has been used in tests of temporal discrimination in rats (Church et al., 1991), with no reported aversive qualities, and likewise for tones (delivered by Sonalert apparatus) delivered at higher intensities, such as 93dB (M.S. Matell, personal communcation).

could be before the rats switched their preferential choice back to the smaller food amount.[33]

So, in this modifed design, a difference in the value between the reinforcement schedules is maintained, but these now differ quantitatively as opposed to qualitatively: this reduces the risk of a change in the utility of the unconditional stimuli confounding the results, and also facilitates a 'two-choice' design: i.e. one in which subjects are presented with a number of discrete options to 'record their choice'. In doing so, we bring the design closer to our 'modified ballot paper' example, above, although it's important to note a few important differences. By employing a timed cut-off point for responding[34], a 'two-choice' design could allow for two different, active responses (e.g. press left lever, or press right lever), and also, separately, a 'no response' option (in which no lever press recorded within the pre-defined timescale). However, in effect, our design dispenses with this, employing a procedure in which subjects can only progress to the next trial if they perform either of the responses (i.e. press either lever). Whilst not a 'forced choice' (the subjects were free to do nothing, although were, perhaps, unlikely to do so, given their sessions could be up to an hour long, and there was no other food available), in our electoral analogy, this would, in effect, be 'compulsory voting'[35]. Whilst we dispense with a 'no response' option (the data from which may actually be of some interest), by 'forcing' the animals to record a choice, we potentially glean more information regarding their judgements (the utility of 'forced-choice' procedures is well-recognised in the human psychological literature: for example, when 'forced' to respond, uncertain participants often perform above chance (e.g. Azzopardi & Cowey, 1997)).

In addition to these design modifications, we employ a larger sample size (n=16), use a repeated-measures design, employ a wider range of probe values, with additional sessions presenting probe stimuli of a different design, and adopt a more extensive, flexible method of statistical analysis. Otherwise, we employ a

[33] Richards et al (1997) found a similar pattern, again with rats, using water as reinforcement.

[34] As Harding et al did (10 seconds).

[35] Another option some social commentators have recently advocated!

treatment very similar to that used by Harding et al, and we also conduct a number of concurrent tests. Our experimental hypotheses remain the same: rats undergoing a treatment designed to induce a negative change in affective state (namely, a series of unpredictable housing events), will be less likely to respond to ambiguous probe stimuli as if judging them to have the better outcome of two pellets of food; in addition, we hypothesise that their presses on the lever associated with the larger quantity of food will be slower than the control group, and *vice versa* for presses on the lever associated with the smaller quantity of food.

Please note that in subsequent sections, we take our lead from others (e.g. Harding et al., 2004; Matheson et al., 2008), and operationalise 'optimism' as an increased probability of pressing the lever associated with two pellets of food (and *vice versa* for 'pessimism').

METHOD

Overview

16 rats were trained to criterion on a two-choice operant discrimination task, with differential food reinforcement: i.e. they were trained with food reinforcement to reliably perform one type of response when presented with one type of stimulus, and to perform a different type of response when presented with a different stimulus; see Figure 2.2 for a schematic summary. Following presentation of a stimulus, the amount of food they received when they performed the corresponding ('correct') response differed between the two stimuli, but was always the same for each; otherwise, they received no food if they performed the response not corresponding to that particular stimulus (the 'incorrect' response). The two responses (the *conditional responses*, or *CR*) were a press on a lever to the left, or a lever to the right of the food delivery bowl; the two stimuli (the *conditional stimuli*, or *CS*) were auditory tones of 2 or 4kHz; and the two different amounts of reinforcement (the *unconditional stimuli*, or *US*) were 1 or 2 pellets of food.

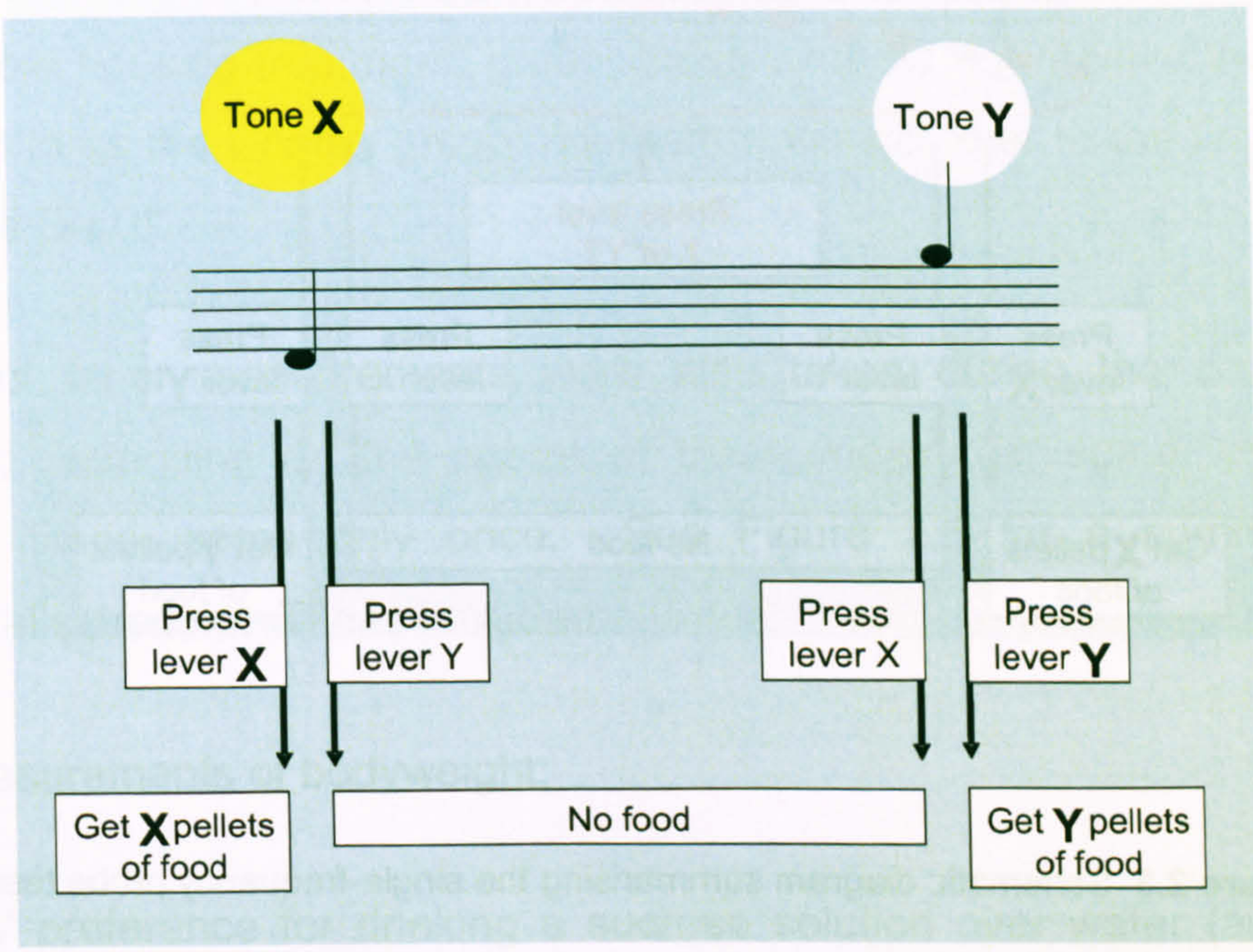

Figure 2.2 Schematic diagram summarising the two-choice operant discrimination task the *subjects* were trained to perform.

Once trained in this task, they were then presented with a series of 'probe' stimuli: namely, auditory tones which differed in some way from those with which they had been trained, and their responses to these, with respect to their choice of lever pressed and their latency to do so, were recorded. Some of these probe stimuli were single tones of a different frequency to those with which they had been trained (e.g. 1.6kHz, 3.2kHz, etc.; hereafter, we refer to these as 'single-frequency' probes; Figure 2.3), whereas others consisted of the two training tones (i.e. 2kHz and 4kHz) played together (we'll refer to these as 'dual-frequency' probes; Figure 2.4).

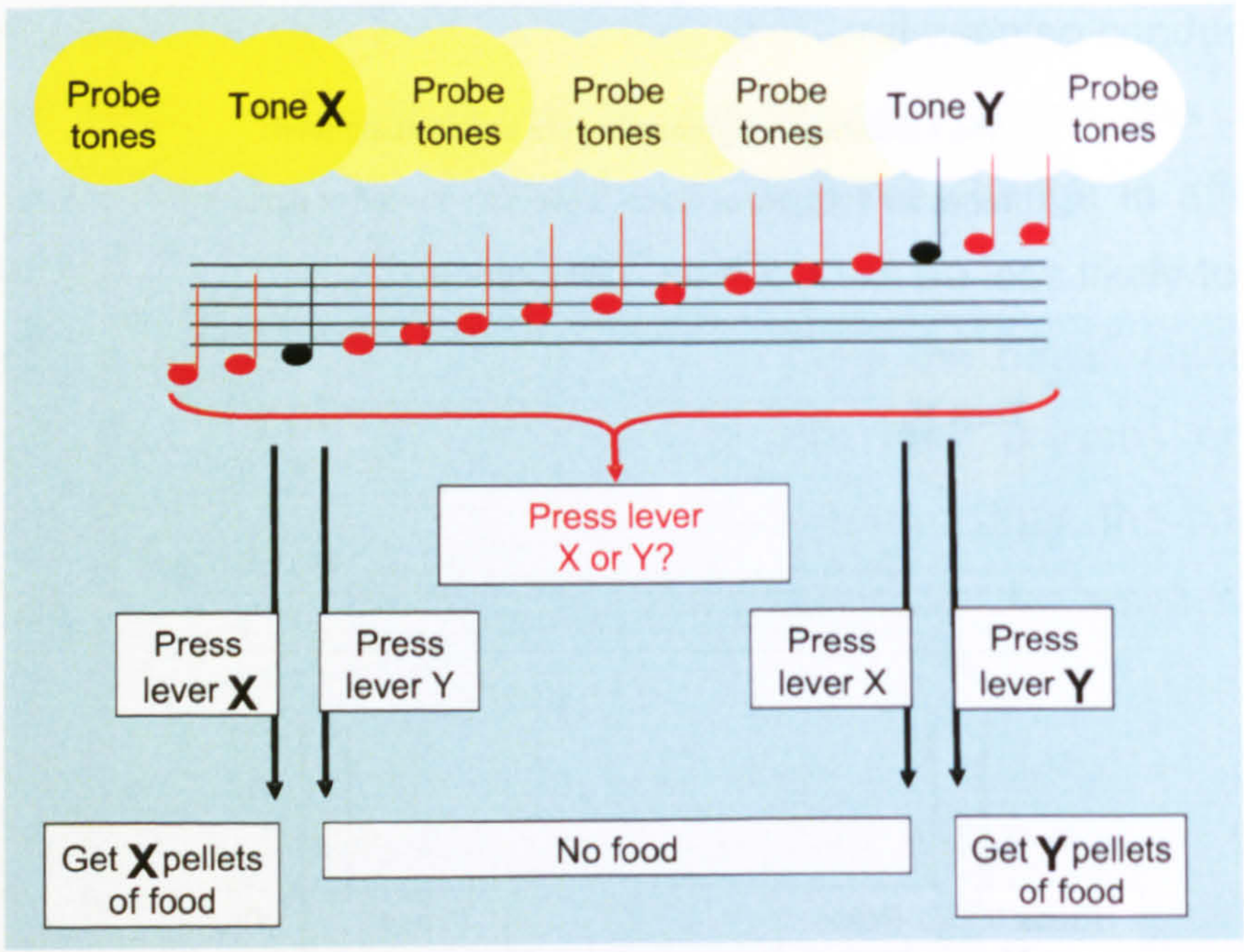

Figure 2.3 Schematic diagram summarising the single-frequency probe test.

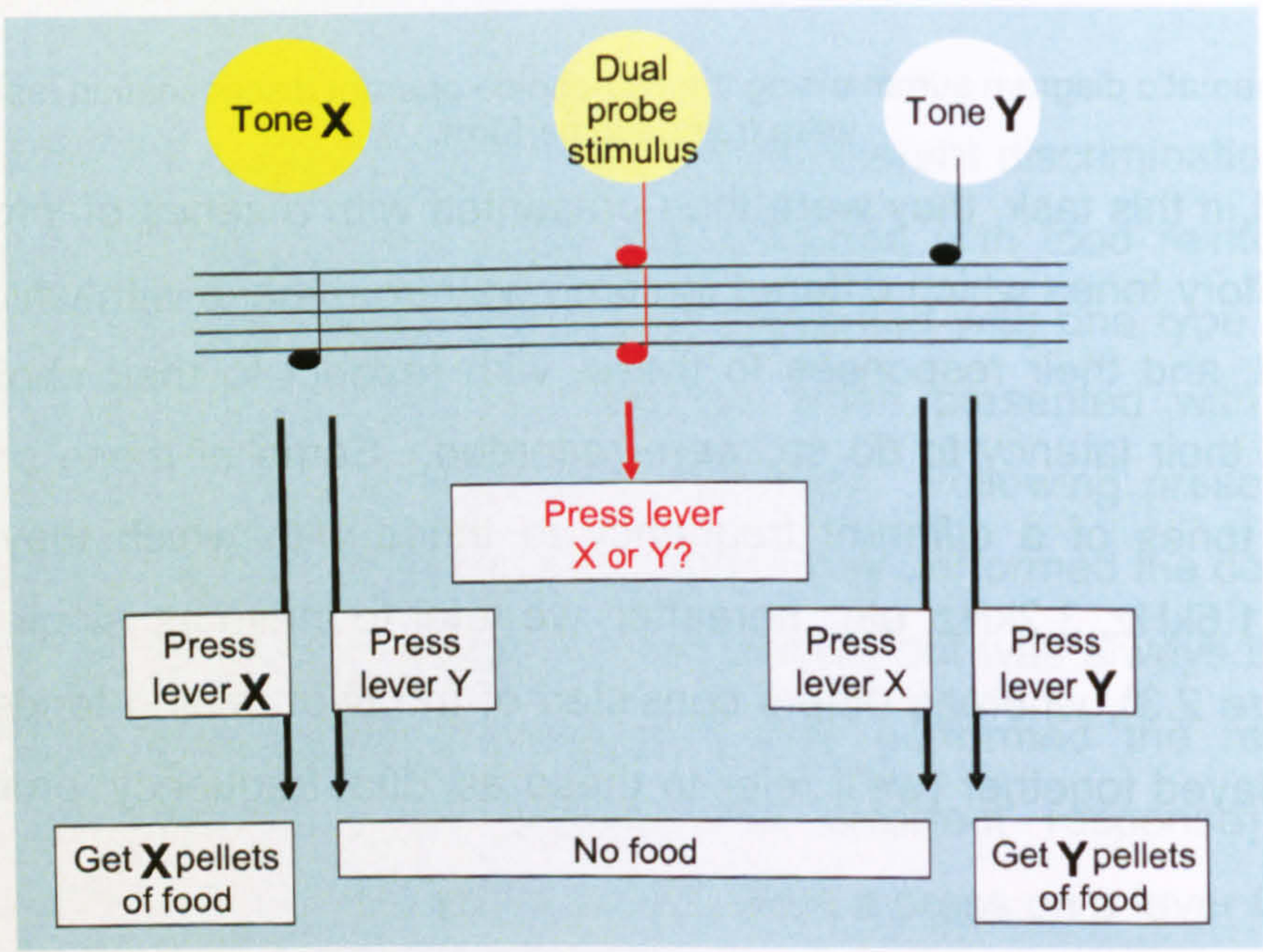

Figure 2.4 Schematic diagram summarising the dual-frequency probe test.

All the subjects were presented with these probe stimuli on two occasions: once before (*Phase 1*), and once during (*Phase 2*) a treatment designed to induce a change in their affective state, namely a period of unpredictable housing; half the rats underwent this *treatment,* and are hereafter referred to as the *UHT* (*unpredictable housing treatment*) group, those who did not receive this treatment are referred to as the *Control* group. Hereafter, we will refer to the first day of the *treatment* as Day 1.

A variety of other measurements were also taken during the course of the experiment; depending on the nature of these measures, some were taken a number of times, some only once. See Figure 2.5 for a summary of the experimental schedule. These included:

- measurements of bodyweight;

- rats' preference for drinking a sucrose solution over water (as mentioned earlier, the sucrose preference test was developed as an indicator of anhedonia in rodents (Willner et al., 1987), and has been widely used since (as reviewed in, for example, Willner, 2005));

- the time they took to eat a certain amount of food (this task has previously been employed by Harding et al (2004) as a test of feeding motivation (after Abeyesinghe, 2000));

- the number of times they pressed a lever to receive food, with the number of lever presses required to receive food increasing each time food was delivered (such progressive schedules have been employed elsewhere as indicators of food motivation (e.g. Bokkers et al., 2004; Ferguson & Paule, 1997; Schutz et al., 2006)).

- their behaviour on an 'elevated plus maze' (widely used as a measure of anxiety (for reviews see, for example: Carobrez & Bertoglio, 2005; Walf & Frye, 2007));

- their behaviour in an 'open field' arena, and their behaviour in the 'open field' arena once a novel object had been introduced (the open-field test, and the novel object test, are often used as measures of anxiety-like behaviour (e.g. Heisler et al., 1998; Prut & Belzung, 2003; van Gaalen & Steckler, 2000)).

The training stimuli (2kHz or 4kHz), responses (left or right lever), quantity of food reinforcement (1 or 2 pellets), experimental room in which the operant training and testing took place, and *treatment* group were all counterbalanced. When we later make reference to *contingency* groups, this refers to the counterbalancing of training stimuli (2kHz or 4kHz) with quantity of food reinforcement (1 or 2 pellets): those in which reinforcement with 2 pellets of food was associated with the 2kHz tone (the *2kHz=2pell* group), and those in which it was associated with the 4kHz tone (the *4kHz=2pell* group).

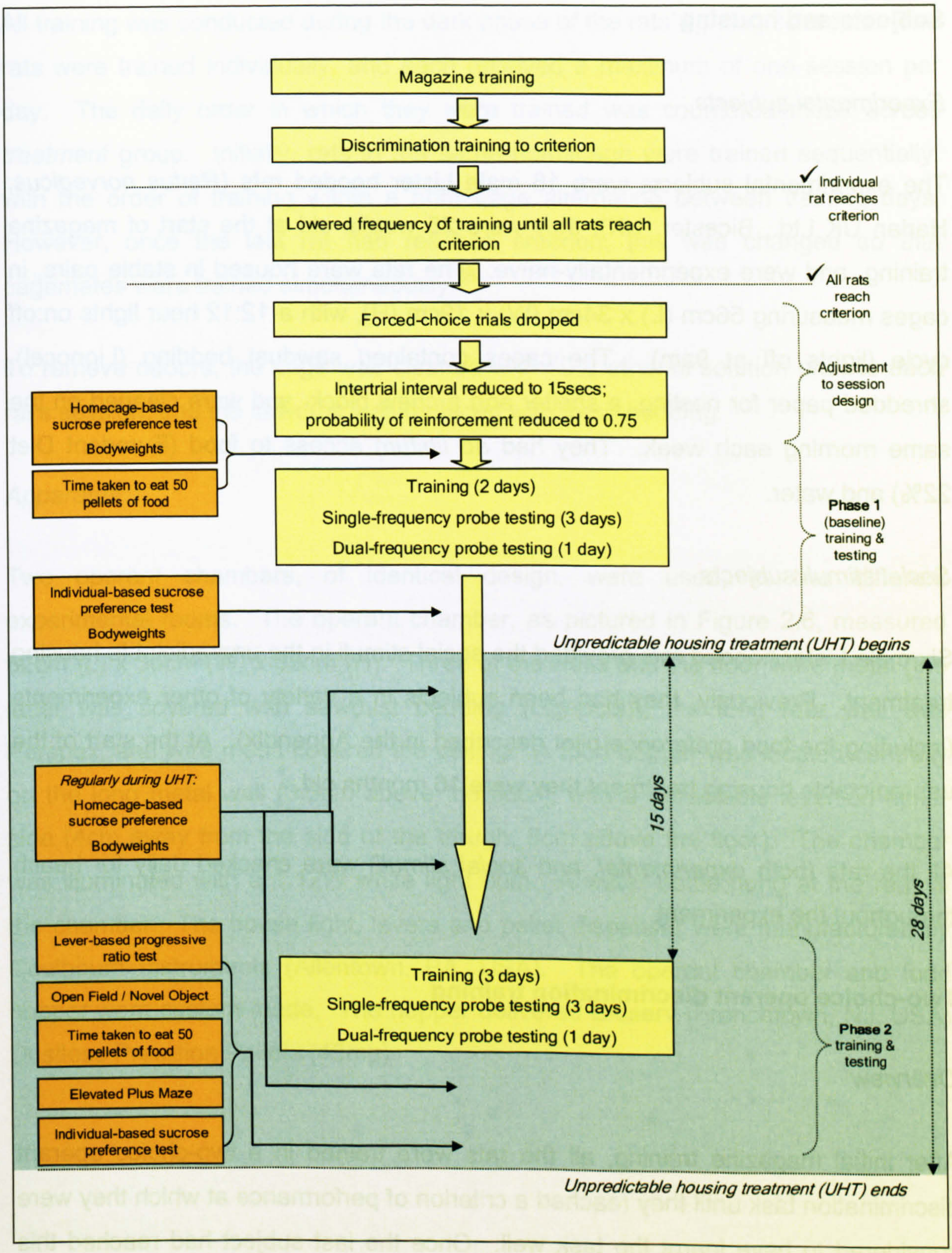

Figure 2.5 A summary of the experimental schedule.

Subjects and housing

Experimental subjects

The experimental subjects were 16 male Lister hooded rats (*Rattus norvegicus;* Harlan UK Ltd., Bicester, UK); they were 25 weeks old at the start of magazine training, and were experimentally-naïve. The rats were housed in stable pairs, in cages measuring 56cm (L) x 34cm (W) x 19cm (H), with a 12:12 hour lights on:off cycle (lights off at 9am). The cages contained sawdust bedding (Lignocel), shredded paper for nesting, a shelter and a chew block, and were cleaned on the same morning each week. They had *ad libitum* access to food (Eurodent Diet 22%) and water.

Social stimuli subjects

Six male Lister hooded rats provided the social stimuli in the unpredictable housing treatment. Previously, they had been subjects in a variety of other experiments (including the food preference pilot described in the Appendix). At the start of the unpredictable housing treatment they were 16 months old.

All the rats (both *experimental*, and *social stimuli*) were checked daily for health throughout the experiment.

Two-choice operant discrimination training

Overview

After initial magazine training, all the rats were trained in a two-choice operant discrimination task until they reached a criterion of performance at which they were considered to have learnt the task well. Once the last subject had reached this criterion, aspects of the task design were changed so that it more closely resembled that of the probe testing sessions (M. Bateson, personal communication): once the rats had undergone a number of additional sessions with this modified design, they underwent probe testing.

All training was conducted during the dark phase of the rats' lighting schedule. The rats were trained individually, and each received a maximum of one session per day. The daily order in which they were trained was counterbalanced across *treatment* group. Initially, rats in the same homecage were trained sequentially, with the order of training within a homecage alternating between training days. However, once the last rat had reached criterion, this was changed so that cagemates were trained simultaneously.[36]

To remove odours, the cage was cleaned with 70% ethanol solution prior to each rat's session, and the sawdust was replaced with fresh bedding.

Apparatus

Two operant chambers, of identical design, were used, in two different experimental rooms. The operant chamber, as pictured in Figure 2.6, measured 52cm (L) x 30cm (W) x 35cm (H). Three of the walls and the floor were metal (the latter was covered with sawdust bedding (Lignocel)), the long rear wall was Perspex, and wire mesh covered the ceiling. A food hopper was located centrally on the long metal wall (3.5cm above the floor), with a retractable lever on either side (4cm away from the side of the trough, 8cm above the floor). The chamber was illuminated with a 1.12W white light bulb. A water bottle hung at the rear of the chamber. The house light, levers and pellet dispenser were manufactured by Coulbourn Instruments (Allentown, PA, USA). The operant chamber and food hopper were custom-made. The hopper delivered Bioserv (Frenchtown, NJ, USA) Dustless Precision Pellets (45mg).

[36] Since the pre-existing schedule meant some rats received a period of solitary housing prior to their operant session, which may have induced a change in affective state, this revised schedule was preferred.

Figure 2.6 The operant chamber, minus the wire mesh ceiling (and the speaker(s), house light, and water bottle attached to it). The retractable levers can be seen either side of the food hopper, which dispensed food pellets stored in the carousel above.

For all training sessions and single-frequency probe sessions, a single speaker (Coulbourn Instruments) was placed centrally, at ceiling level above the food hopper, facing down into the chamber. For all dual-frequency probe sessions two speakers were placed in approximately the same location, adjacent to each other on the midline of the chamber. Tones were produced by a programmable tone generator (A12-33, Coulbourn Instruments): the Hertz frequency output of these units was calibrated to 1% accuracy using a Vision 8-Channel Data Acquisition System (LDS Test & Measurement Ltd, Herts., UK). The volume of the speakers was adjusted so the tonal intensity was 70dB at the approximate location of a rat's ears when he was sitting between the levers, in front of the hopper.

The house light, levers, pellet dispensers, tone generators and speakers were operated by Graphic State (v3.02) software (Coulbourn Instruments).

Magazine-training

All rats received 6 sessions of magazine training. Each session started with the presentation of one lever[37] which, if pressed, resulted in the immediate delivery of one food pellet. If 10 lever presses were made, the lever was retracted, and the other lever was presented and reinforced on the same schedule. This alternating pattern continued until the rat had made 60 lever presses, or 60 minutes had elapsed, whichever came first. In addition, regardless of any lever-pressing, a parallel schedule was in operation in which the active lever retracted for one second prior to the automatic delivery of one food pellet. For the first three sessions, this autoshaping procedure occurred every minute, for the final three sessions it occurred every 3 minutes. This magazine training procedure was based on Mattel & Meck (1999).

In the final session of magazine training, all rats pressed the levers 60 times, and each ate all the pellets dispensed (bar two rats who left one each).

Two-choice operant discrimination training, with differential reinforcement

In a training session lasting 40 minutes, each rat was presented with a series of trials consisting of a two-second presentation of an auditory tone (either 2kHz or 4kHz), followed by the presentation of both (free-choice) or one (forced-choice) lever(s); as soon as the rat pressed a lever (assuming it did before the session terminated), the lever(s) were retracted, and food was delivered as appropriate. By default, the trials were free-choice, and each auditory tone (2kHz or 4kHz) had an equal probability of being presented. If the rat pressed the 'correct' lever (i.e. the one designated by the experimenter as corresponding to that stimulus), either 1 or 2 pellets of food were delivered, depending on the identity of the auditory tone. If the rat pressed the 'incorrect' lever, the following trial was forced-choice: i.e. the same auditory tone was played again, and only the 'correct' lever was presented. The inter-trial interval (ITI) was 30 seconds.

---

[37] Either left or right; for each rat, the position of the first lever presented alternated between sessions, and the position of the first lever presented in the first session was counterbalanced with respect to future group assignment.

The rats received one session per day, with an occasional day off, until they reached criterion. Criterion was defined as three consecutive sessions in which performance in free-choice trials was significantly greater than chance, as judged by a binomial test, for each trial type (e.g. Harding et al., 2004). Once a rat had reached this criterion, it received two further sessions over the next two training days, before the amount of training it received was reduced to a lower rate (once or twice per week, depending on its performance[38]) until the last rat had reached criterion[39].

Once all the rats had reached criterion, the design of the operant task was changed so that it more closely resembled the probe sessions: namely, the forced-choice trials were dropped, the probability of reinforcement following a correct response was reduced to 0.75, and the ITI was decreased to 15 seconds. To guard against long sequences (i.e. runs) of the same trial type, the selection of trials was pseudorandomised, into blocks of eight trials. Each block consisted of three reinforced trials for each tone, and one non-reinforced trial for each tone. Trials were selected, at random, from this list, without replacement; when the last trial had been selected, selection began from the next block, which was of exactly the same composition as the last.

The rats underwent a few more sessions with this modified task[40] before probe-testing commenced, including on each of the two consecutive days before probe-testing in *Phase 1* (i.e. at 'baseline'), and over the three consecutive days before probe-testing began in *Phase 2* (Days 16–18 of the *treatment*). See Figure 2.7 for a summary of the final training protocol.

---

[38] Each rat received one session per week, unless it failed to perform above chance (as judged by a binomial test) for both auditory tones in a given session, or for one tone only for two sessions in a row, in which case the frequency increased to twice per week, until the rat performed above chance for both auditory tones, at which point the reverted back to frequency of training sessions reverted back to once per week.

[39] Since the unpredictable events in the unpredictable housing treatment (UHT) took place in cages other than the rats' homecage, it was thought best to start the treatment after all the rats had reached criterion: otherwise, if a rat who was still being trained was pair-housed with a rat in the *UHT* group who had reached criterion, and whose treatment had therefore commenced, the former rat would have prolonged periods of pre-treatment solitary-housing whilst the latter underwent the unpredictable events.

[40] Again, the number of sessions a rat received depended on its performance; unfortunately, due to problems with the building in which the rats were housed, we couldn't immediately commence the probe-testing and treatment phase.

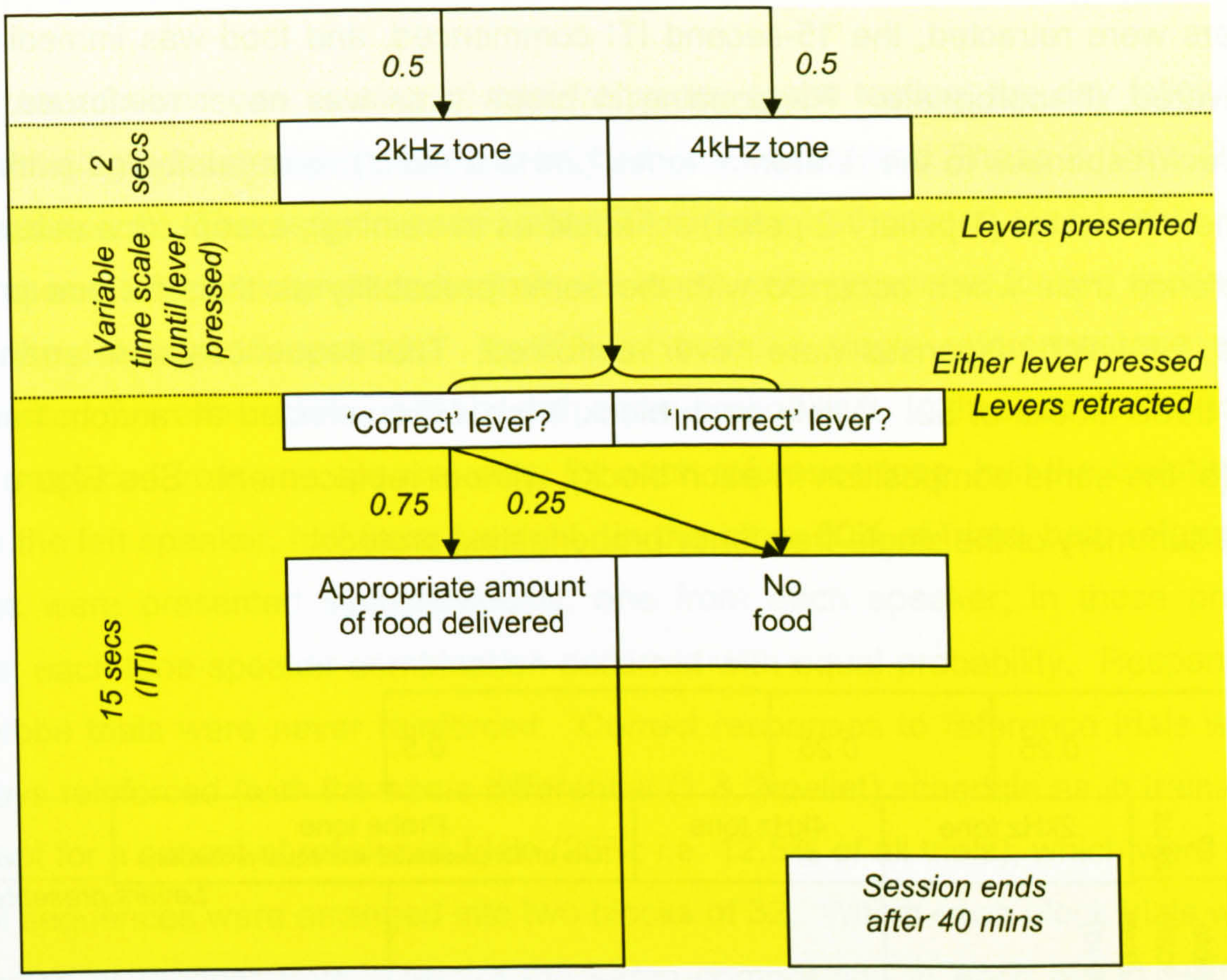

Figure 2.7 Flow chart summarising the final two-choice operant discrimination training protocol. The probabilities stated here refer to the session as a whole.

Single-frequency probe testing

The rats received three sessions of single-frequency probe testing, on three consecutive days, both at *Phase 1* (i.e. at 'baseline', before the *treatment* had begun) and at *Phase 2* (Days 19-21 of the *treatment*). Sessions terminated after 156 trials, or 60 minutes had passed, whichever came first. All tones were presented for two seconds. In 50% of the trials, the reference tones were presented (2kHz & 4kHz, with equal probability). Otherwise, one of 13 probe tones was presented (again, with equal probability). The probe tones ran from 1.6kHz to 4.4kHz in 200Hz increments (not including the reference frequencies), meaning there were nine probe tones of a frequency intermediate to the two reference tones, and two at either far end. All trials were free-choice (i.e. both levers were

presented following the 2-second tone); once a response had been made, both levers were retracted, the 15-second ITI commenced, and food was immediately delivered, if appropriate. Responding in probe trials was never reinforced. All correct responses to the reference tones (2kHz & 4kHz) were reinforced (with the same differential (1 pellet / 2 pellet) schedule as in training), except for a subset of reference trials which occurred with the same probability as trials for one probe tone, for which responses were never reinforced. Trial sequences were arranged into three blocks of 52. Within each block, trials were selected at random from a list (of the same composition in each block), without replacement. See Figure 2.8 for a summary of the single-frequency probe testing protocol.

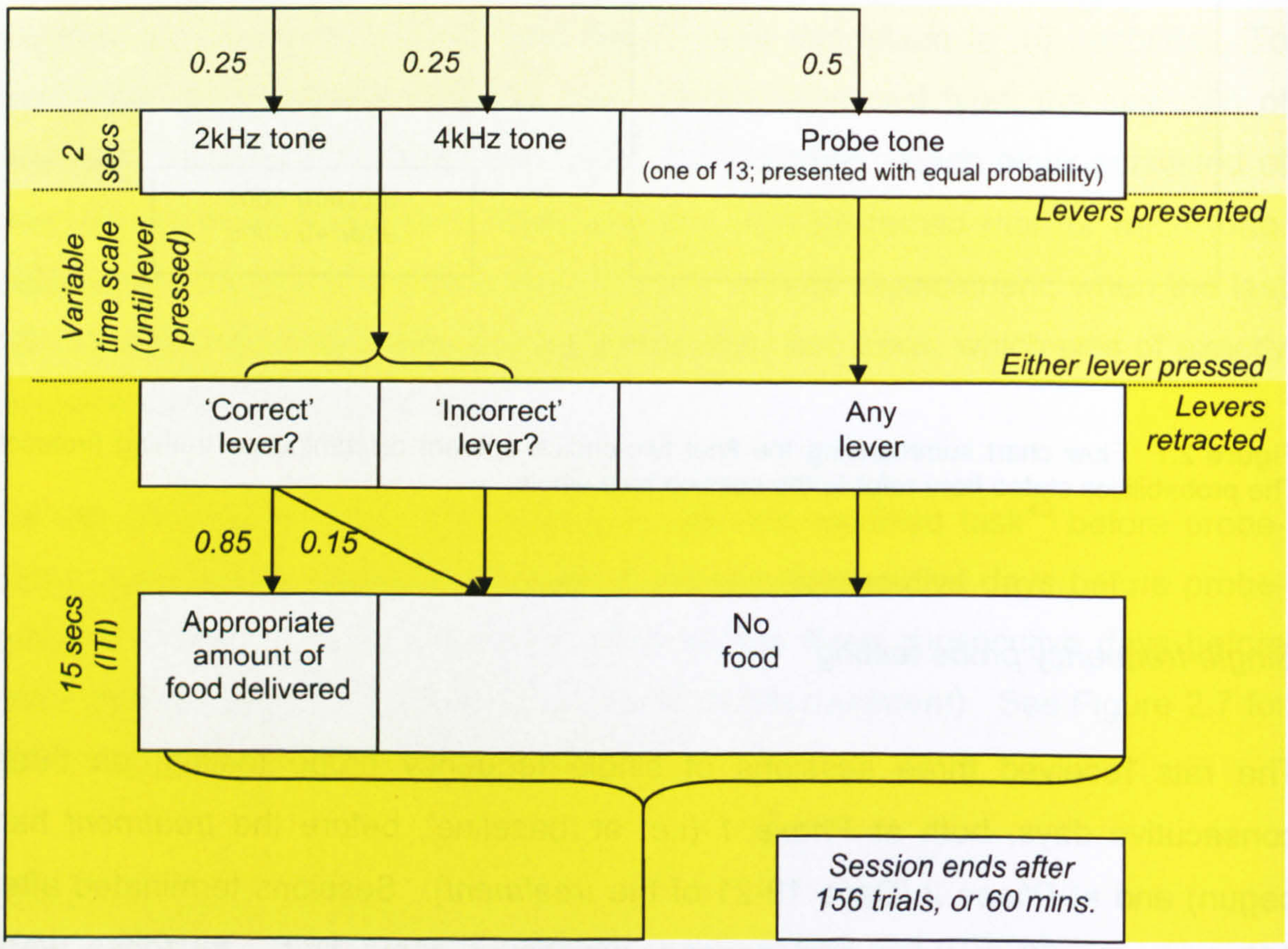

Figure 2.8 As Figure 2.7, but summarising the protocol for the single-frequency probe test sessions.

Dual-frequency probe testing

The rats received one session of dual-frequency probe testing, the day following the last single-frequency probe session, both in *Phase 1* and *Phase 2* (Day 22 of the *treatment*). The session ended after 30 minutes had passed, or 64 trials had been completed, whichever came first. All tones were presented for two seconds, and the ITI was 15 seconds. In 50% of trials, a single reference tone was presented (i.e. either 2kHz or 4kHz, with equal probability). In these trials, the tone was produced from one speaker only; for each reference tone, half the trials were from the left speaker, half from the right. In the other 50% of trials, both reference tones were presented simultaneously, one from each speaker; in these probe trials, each tone-speaker combination occurred with equal probability. Responses to probe trials were never reinforced. Correct responses to reference trials were always reinforced (with the same differential (1 & 2 pellet) schedule as in training), except for a subset of reference trials (25%; i.e. 12.5% of all trials), which were not. Trial sequences were arranged into two blocks of 32. Within each block trials were selected at random from a list (of the same composition in each block), without replacement. See Figure 2.9 for a summary of the dual-frequency probe testing protocol.

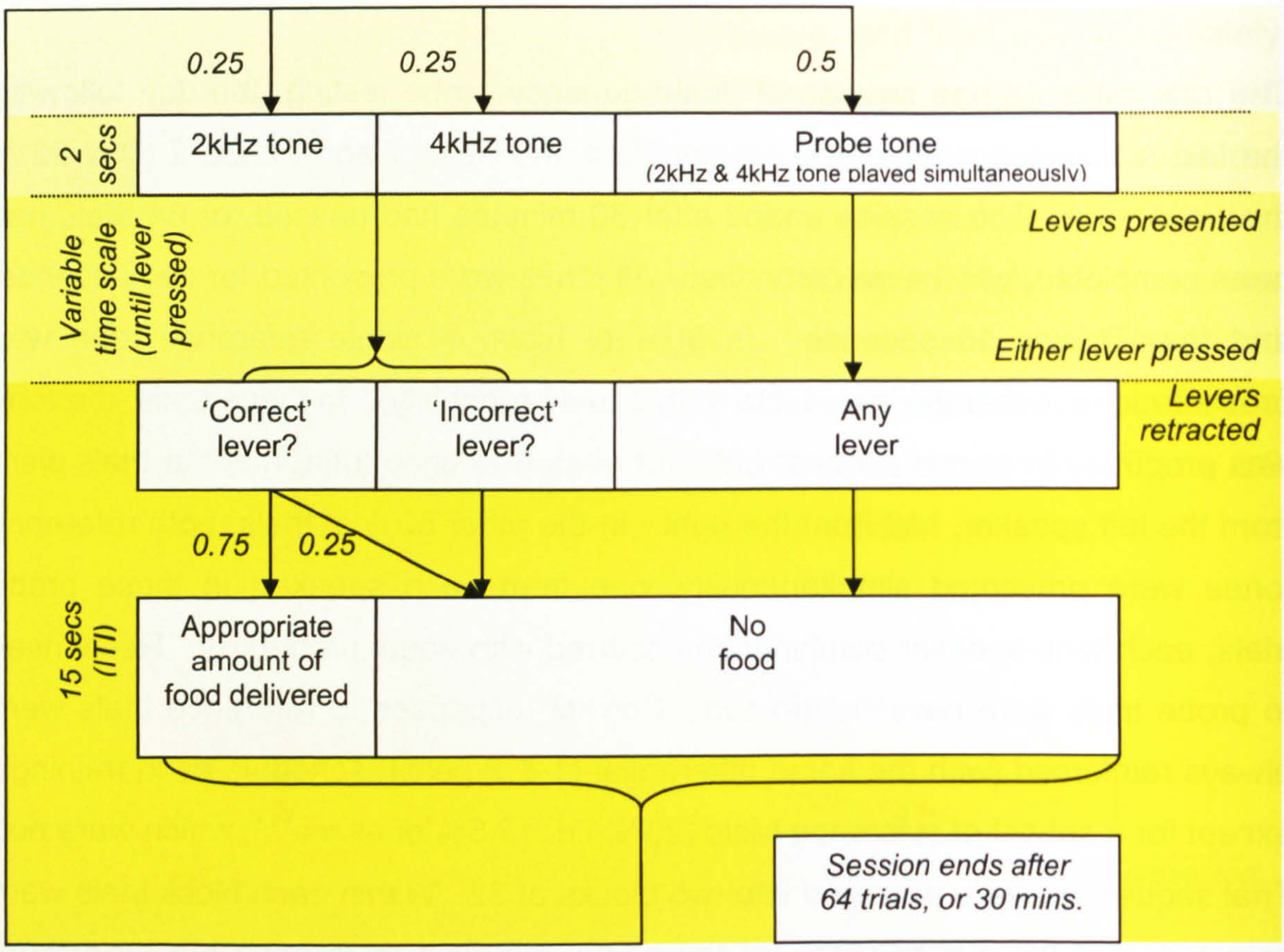

Figure 2.9 As Figure 2.7, but summarising the protocol for the dual-frequency probe test sessions.

Concurrent tests

All concurrent tests were conducted in the dark phase of the rats' lighting schedule (see Figure 2.5 for a summary of their timetabling).

Bodyweight

The rats were weighed, in counterbalanced order, once before, three times during (on Days 7, 14 & 21), and once after the *treatment* (on the day following its termination).

Individual sucrose preference test

Each rat was placed in a test cage with Lignocel bedding, shredded paper, *ad libitum* food (Eurodent Diet 22%), and two drink bottles: one containing water, the other 1% sucrose solution. Each filled drink bottle was weighed just prior to the test, and their starting position (i.e. left or right-hand side of the cage) was counterbalanced across treatment. After 90 minutes, the drink bottles were weighed and placed back in the test cage in a reversed position. The bottles were again weighed after another 90 minutes, when the test ended. These tests were conducted twice: once before the start[41], and once at the end (Day 28) of the *treatment* phase.

Homecage-based sucrose preference test

The *Individual Sucrose Preference Test,* described above, involved moving all rats into test cages, and it is possible that the resulting novelty and isolation may have induced a negative change in affect; i.e. it may have been a stressful event. Since we wished to conduct sucrose preference tests with all *subjects* during the *treatment* phase, and wished to minimise any such disruption to the *Control* group, we therefore also conducted sucrose preference tests in the rats' homecages.

The rats remained in their pair-housed homecages, which were unchanged except for their water bottles being removed at the start of the test and replaced with two drink bottles, one containing water, the other 1% sucrose solution. The procedure thereafter was the same as for the *Individual* test described above, except that the bottles were weighed and reversed after 4 hours, with the test lasting a total of 8 hours.

[41] Following probe-testing conducted during *measurement phase 1.*

These tests were conducted once before the start[42], on two occasions during (Days 7 & 14), and once at the end, of the *treatment* phase (on the day following its termination).

Time taken to eat 50 pellets of food

Rats were tested individually, in their homecage; their cagemate was removed to a holding cage during the test. The shelter and chewblock were removed, and a brown bowl containing 50 food pellets (as used in their operant training and testing) was placed in the centre of the cage (the bowl was washed with 70% ethanol solution, then dried, prior to each trial). The time from when they first took a pellet into their mouth until all were eaten was recorded. This test rat was then swapped with his cagemate in the holding cage, who was then tested in an identical manner. At the end of the test, the shelter and chewblock were replaced, and both rats were returned to their homecage. Order of testing was counterbalanced across treatment group. These tests were conducted twice: once before the start[43], and once at the end, of the *treatment* (on the day following its termination).

Lever-based progressive ratio test with food reinforcement

Rats were tested individually, in the operant chamber used for their discrimination training and probe-testing. The session started with the '2-pellet' lever (i.e. the lever on which correct responses in their operant training and test sessions had been reinforced with 2 pellets of food) being presented. The first lever press resulted in the immediate delivery of 2 pellets, and thereafter presses were reinforced on a progressive ratio of 5 (PR5), starting with 5 presses (i.e. 2 pellets were delivered after 5, then a further 10 (total presses: 15), then a further 15 (total presses: 30), etc. presses). If no reinforcement took place for 5 minutes, or 60 minutes elapsed, the session terminated.

---

[42] Prior to probe-testing conducted during *measurement phase 1.*

[43] Following probe-testing conducted during *measurement phase 1.*

The test was administered once, towards the end (Day 24) of the *treatment* phase. One of the *subjects* in the *Control* group was excluded from this task due to ill-health on the test day.

Elevated plus maze (EPM)

The elevated plus maze (Coulbourn Instruments) was raised 55cm from the floor, and consisted of 4 arms at right angles to each other, connected by a central hub allowing the rat to move between them. Each arm was 50cm long, and 10cm wide. The two 'closed' arms were on opposite sides of the hub, and had vertical walls, 30cm high, either side of the runway (but not at the terminal end). The other two arms were 'open', and had no walls, although they did have a small raised edge. The maze was made of black, opaque Perspex, and was evenly-illuminated (to our eyes) by two 60W red lights directed towards the ceiling corners of the experimental room.

The walls and runways were sprayed with 70% ethanol solution, and then wiped dry, prior to each rat's test. Rats were tested individually, in a counterbalanced order. Each rat was placed in the centre, facing the same closed arm, and was filmed for 5 minutes, at which point the test ended.

The video recordings were analysed by a volunteer who was otherwise not involved with the experiment (except for analysing the video data in the *Open Field Test with Novel Object,* described below) and who was blind to treatment assignment, using Observer (v5.0) software (Noldus, Wageningen, The Netherlands). The rat's location in the arena (central hub, each open arm, each closed arm) was recorded, with the rat judged to have entered an area when all four paws were in it. From these observations, the *total number of crossings across area boundaries,* the *percentage of test session time spent in the open arms,* and the *latency to first enter an open arm*, were calculated. In addition, the percentage of test session time spent performing each of the following behaviours,

together with the number of discrete bouts of each, was recorded: *grooming*, *rearing*[44], and *head dipping from the open arms*[45].

The test was administered once, at the end (Day 28) of the *treatment* phase.

Open field test, with a novel object

The open field was a square, Perspex arena, with white, opaque walls and floor, and a transparent roof. The walls were 30cm high, and 65cm long. A transparent plastic food container filled with sand constituted the novel object. It was square, with a sealed lid, and measured 6cm high, and 11cm long. The arena was evenly-illuminated (to our eyes) by a 60W red light positioned centrally above it, directed towards the experimental room's ceiling.

Both the arena and novel object were sprayed with 70% ethanol solution, and wiped dry, before each rat's session. Rats were tested individually, in a counterbalanced order. Each rat was placed into the empty arena in the same corner, facing the centre, and then filmed for 15 minutes. After the first 10 minutes, the novel object was placed in the centre of the arena. The first 5 minutes constituted the open field test, the final 5 minutes the novel object test (i.e. an object of relative novelty was placed in an environment of increasing familiarity; e.g. Francia et al., 2006; Popovic et al., 2004; Powell et al., 2004).

The video recordings were analysed by the same volunteer who analysed the EPM, again blind to treatment assignment, using Observer (v5.0) software.

For the open field test, the rat's location in the arena (divided into nine squares of equal area: one central and eight peripheral[46]) was recorded, with the rat judged to have entered an area when all four paws were in it. From these observations, the

---

[44] Both 'supported' rearing, with forepaws on the maze walls, and 'unsupported' rearing, with forepaws not touching the walls.

[45] Defined as exploratory movement of head and/or shoulders over the edge of the maze.

[46] This was done *post-hoc*, i.e. traced on the screen showing the footage from the test, rather than marked out on the test arena floor itself.

total number of areas entered, and the *percentage of test session time spent in the central area of the arena*, was calculated. In addition, the percentage of test session time spent *grooming,* and the percentage of test session time spent *rearing*[47], were also recorded, together with the number of discrete bouts of each.

For the novel object test, the arena was divided differently, into a central square (43cm long, containing the novel object with an 11cm boundary of floor space surrounding it) and the remaining peripheral area. The rat's location during the test session, with respect to these two areas, was again recorded, using a 4-paw criterion. From these observations, the *percentage of test session time spent in the peripheral area* (i.e. away from the novel object) was calculated.

These tests were administered once, at the end (Day 27) of the *treatment* phase.

Unpredictable housing treatment

After the completion of all *phase 1* tests, half the rats underwent unpredictable changes in their husbandry regime designed to be mildly stressful (the *unpredictable housing treatment*, or *UHT.* This procedure was devised by Harding et al (2004) to model stressful events which might occur as part of a (very negligent) lab husbandry system (as mentioned earlier, this was, in turn, adapted from the chronic mild stress procedure; Willner, 2005; Willner et al., 1987).

The treatment lasted for 28 days, and consisted of five different husbandry events which could occur at any time during the dark phase of the rats' lighting schedule. These events, and the maximum frequency with which they could occur, are summarised in Table 2.1. No more than two events occurred on any given day, and they did not overlap. All *UHT* rats underwent these events individually, in a clean test cage, except for the *unfamiliar homecage* event, which took place in an unfamiliar conspecific's homecage: in this instance, the cage was monitored, and

[47] Both 'supported' rearing, with forepaws on the maze walls, and 'unsupported' rearing, with forepaws not touching the walls.

the event terminated if damaging aggression was seen to occur[48]. In the *unfamiliar odour* event, the cage had been vacated in the preceding 15 minutes by an unfamiliar conspecific resident in there for at least one hour. During the *treatment* phase, *UHT* events did not occur during an interval of two hours duration preceding behavioural testing/training, nor during an interval of two hours following the end of such training/testing.

Unpredictable Housing Event	Max. duration (hours)	Max. frequency (per week)
Wet bedding: Bedding dampened with 100ml water.	12	1
Unfamiliar odour: Placed in cage containing unfamiliar conspecific's odours.	4	3
Cage tilt: Tilted by 30°.	7	2
Unfamiliar homecage: Placed in homecage of unfamiliar conspecific (whilst present).	2	3
Light cycle reversal Taken from dark to light environment.	2	3

Table 2.1 Summary of events in the *unpredictable housing treatment (UHT)*.

Data analysis

Overview

To aid interpretation of some of the analyses, and to better characterise the psychophysical properties of the stimuli, the single-frequency probe values were converted into a standardised scale; we describe this below.

Furthermore, the *lever choice* and *latency* data from the single-frequency probe sessions were analysed using multilevel multiple regression models using MLwiN

[48] A number of aggressive encounters were seen to occur: these were generally quickly resolved, and appeared not to be physically damaging. However, on two occasions, the treatment *was* terminated due to concerns for the rats' physical welfare as a result of aggressive interactions taking place; on one of these occasions, it was the experimental subject who was judged to be the main aggressor, whilst on the other occasion it was the social stimulus rat.

2.02 (Rasbash et al., 2005). This was subsequent to exploratory analyses using repeated-measures ANOVAs (using SPSS 14.0), described in Appendix C. We do not present both types of analysis in the Results section, for the sake of brevity, but it is good practice, when developing such multilevel models anew, to cross-check with alternative analytical approaches, for example to guard against errors in model-specification (e.g. Rasbash et al., 2005). This will ultimately allow us to use the more complex (but potentially more informative and flexible) multilevel analyses with some confidence, both in this, and subsequent, chapters. Thus, we refer to the corresponding repeated-measures ANOVAS when summarising our multilevel analyses in the Results section. Below, we introduce our multilevel modelling procedure.

All other analyses were conducted in SPSS 14.0, and all met the assumptions of the statistical procedures used, except where we highlight an issue. When conducting repeated-measures ANOVAs with a within-subjects factor which had more than 2 levels (i.e. $k>2$), we follow the advice of Quinn & Keough (2002), and reject the null hypothesis if either the adjusted univariate output, or the multivariate output, reports significance at the 0.05 level.[49] We present a summary of our results at the end of each subsection.

Converting the single-frequency probe values into a standardised scale

To aid analysis, the tonal frequency values of the stimuli used in the single-frequency probe test sessions (ranging from 1.6 – 4.4kHz, at 0.2kHz intervals) were log-transformed onto a scale which was standardised around the quantity of the associated food reinforcer, so that the values of "1" and "2" on the resulting scale corresponded to the tonal frequencies (2kHz or 4kHz, depending on the counterbalanced *contingency* group) associated with one and two pellets of food, respectively. The scale was designed so that the intervals between probe and reference stimuli increased in magnitude as the tonal frequencies (kHz) to which

[49] The output in SPSS provides a formal test of sphericity (Mauchly's), but this is not reliable when the assumption of multivariate normality is not met, and therefore Quinn & Keough follow others' recommendations in routinely inspecting, and reporting, only the output which does not assume sphericity. When the between-subjects factor has only 2 levels (i.e. $k=2$), sphericity is not an issue as the assumption is always met, although it is still necessary to check for homogeneity of variance, and we do so (e.g. Field, 2000).

they correspond decreased in value, resulting in a closer fit to the likely psychophysical character of the auditory stimuli used[50]. See Table 2.2 and Figure 2.10 for numerical and graphical representations of this scale, respectively, and Appendix B (p.316) for details of how the scale was derived.

Original kHz	4kHz = 2 pellets group	2kHz = 2 pellets group
1.6	*0.68*	*2.32*
1.8	*0.85*	*2.15*
2.0	*1.00*	*2.00*
2.2	*1.14*	*1.86*
2.4	*1.26*	*1.74*
2.6	*1.38*	*1.62*
2.8	*1.49*	*1.51*
3.0	*1.58*	*1.42*
3.2	*1.68*	*1.32*
3.4	*1.77*	*1.23*
3.6	*1.85*	*1.15*
3.8	*1.93*	*1.07*
4.0	*2.00*	*1.00*
4.2	*2.07*	*0.93*
4.4	*2.14*	*0.86*

Table 2.2 Log-transformed scales (in *italics*), standardised around reinforcer value, for each of the tonal frequency / food quantity contingency groups.

[50] The perceived difference between two auditory tones will likely depend on the absolute value of those stimuli. For example, the difference between 1.6kHz and 1.8kHz will be perceived as being of a different magnitude compared to the difference between 4.2kHz and 4.4kHz, even though each pair of stimuli are separated from each other by the same difference in Hertz (e.g. Moore, 2004). The just noticeable difference (*jnd;* i.e. the lowest difference in value which can be reliably perceived between stimuli) between two tonal frequencies is a reasonably constant *proportion* (a.k.a. Weber ratio) of absolute tonal frequency across a range of values in rats (Syka et al., 1996; Talwar & Gerstein, 1998); as such, at 4kHz, the *jnd* will be a difference in Hertz value approximately twice as great as that at 2kHz. Therefore, log-transforming tonal frequency provides a scale along which the resulting intervals are likely to be a closer representation of the *perceived* differences between stimuli. Indeed, this is a convention in keeping with the octave scale for pitch, and decibel scale for loudness (e.g. Yost, 2000).

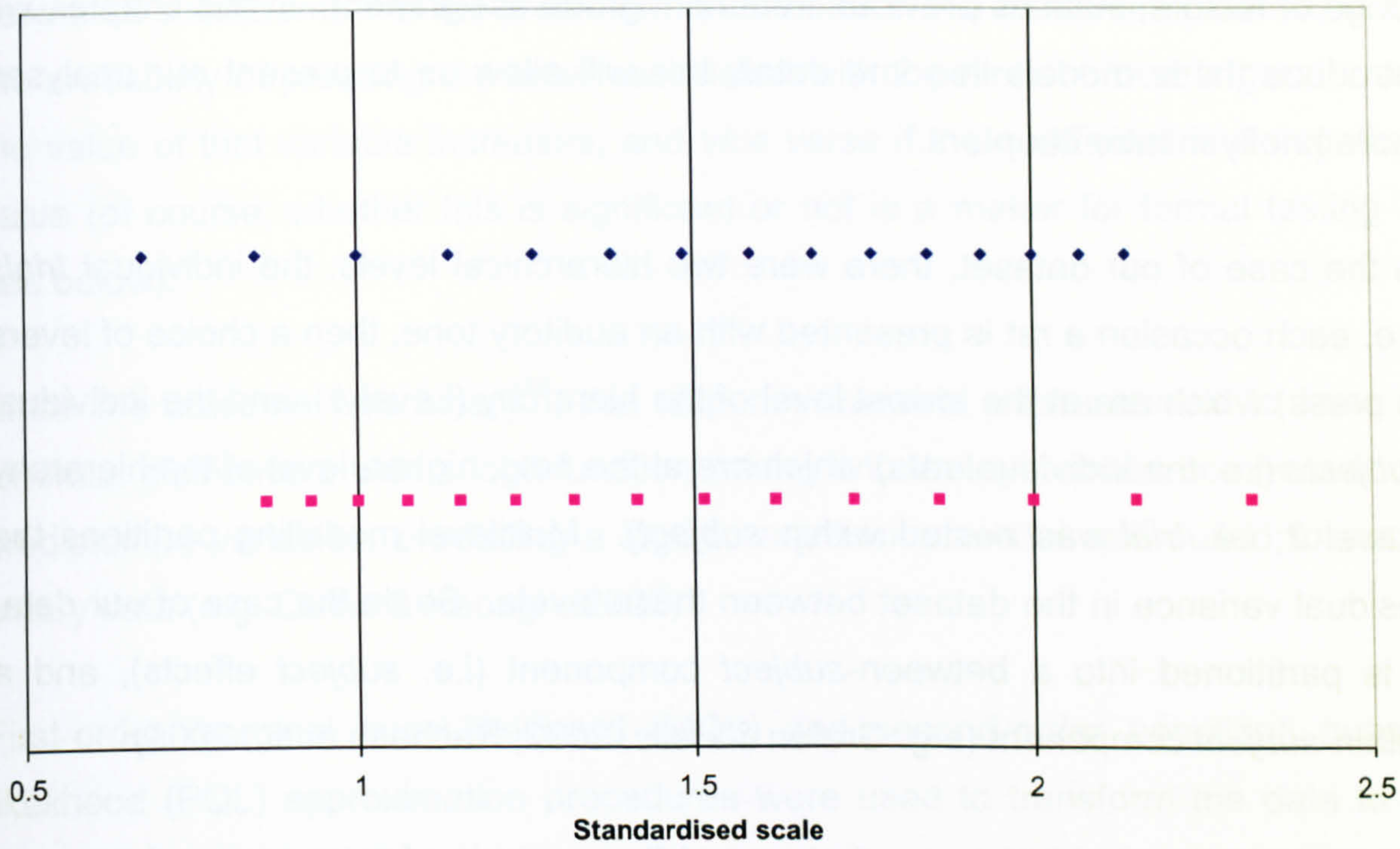

Figure 2.10 Values on a standardised scale corresponding to the tonal frequencies presented to the rats in the single-frequency probe test sessions (blue diamonds: 4kHz = 2 pellets contingency group; purple squares: 2kHz = 2 pellets contingency group).

Multilevel analysis of the single-frequency probe session data: overview

Analyses of the *lever choice* and *latency* data from the single-frequency probe sessions were conducted in MLwiN. For certain analyses, multilevel[51] procedures have benefits over more conventional methods, because they more adequately, and efficiently, model the effects of any hierarchies or clusters in the data. Such effects may be of scientific interest in themselves, but also if they are ignored, or confounded, there is a risk that erroneous conclusions may be drawn (e.g. Goldstein, 2003; Rasbash et al., 2005). More generally, adopting such an approach may yield an analysis which is in some aspects more informative and flexible, allowing us to model a larger dataset, and uncover or clarify effects which may be hidden in an aggregation analysis in which a large number of datapoints are collapsed into a few (e.g. Nezlek, 2001; Quené & van den Bergh, 2004). This approach will also prove useful in subsequent chapters, when we test a wider

[51] Also known as 'random effects models', since at each of the levels, the units (the various *trials* or *subjects*, in the case of our dataset) are conceptualised as being a random sample from a larger population (of trials and subjects, respectively), and it is these larger populations about which the model makes its inferences.

range of factors, such as previous treatment group assignment. In this chapter, we introduce these models in some detail; this will allow us to present our analyses more briefly in later chapters.

In the case of our dataset, there were two hierarchical levels: the individual *trials* (i.e. each occasion a rat is presented with an auditory tone, then a choice of levers to press) which are at the lowest level of the hierarchy (Level 1), and the individual *subjects* (i.e. the individual rats) which are at the next, higher, level of the hierarchy (Level 2; i.e. *trial* was nested within *subject*). Multilevel modelling partitions the residual variance in the dataset between these levels. So, in the case of our data, it is partitioned into a between-*subject* component (i.e. *subject* effects), and a within-*subject* component (e.g. Grafen & Hails, 2002; Rasbash et al., 2005).

For both the analysis of *lever choice*, and the analysis of *latency*, the data were prepared beforehand as follows: only non-reinforced trials[52] were included in the dataset, and any trial in which a lever press was not recorded[53] was excluded. In addition, the continuous predictor (x) variables were centred (e.g. Nezlek, 2001; Rasbash et al., 2005), around their grand mean (i.e. each datapoint was transformed by subtracting the grand mean of all the datapoints in that series from it – see below for a discussion of the merits of centring).

Multilevel analysis of the single-frequency probe session data: lever choice

In the analysis of *lever choice,* the response (y) variable was the choice of lever pressed, with '*1*' indicating that a press was made on the lever associated with 2 pellets of food, and '*0*' indicating no such press was made on that lever (i.e. the lever associated with 1 pellet of food was pressed instead). This means that if the

[52] Whilst most of the reference trials (i.e. in which 2kHz & 4kHz tones were played) presented in the probe sessions were reinforced (i.e. 'correct' responses resulted in food delivery), for a subset of reference trials, food was never delivered, no matter what the rats' response (p.61). These non-reinforced trials appeared as often as each of the probe trials (i.e. in which all tones other than 2kHz & 4kHz were played). Only the non-reinforced reference trials were included in the analysis to maintain a more balanced dataset (with respect to sample sizes, error, and so on).

[53] On the rare occasions (n=3) a session had not finished by the time 60 minutes had elapsed, any unconcluded trial still active at the end of the session was terminated without a lever press being recorded; for all other trials, lever presses were recorded.

estimated coefficients[54] of the predictor (*x*) variables in the model are positive, then the probability of pressing the lever associated with 2 pellets of food is greater as the value of that variable increases, and *vice versa* if the coefficient is a negative value (of course, whether this is significant or not is a matter for formal testing – see below).

Since the outcome was binary[55], the response variable was assumed to have a binomial distribution, with a logit function linking it to the systematic component of predictor (*x*) variables, producing a logistic regression model well suited to such binary data (e.g. Quinn & Keough, 2002).

First-order marginal quasi-likelihood (MQL) and second-order penalized quasi-likelihood (PQL) approximation procedures were used to transform the data to a linear model prior to iterative generalised least squares (IGLS) estimation (e.g. Goldstein, 1986, 2003; Rasbash et al., 2005).

Wald tests were conducted to test the significance of the terms in the model. These tests compare the coefficient estimate with its standard error, and determine the significance of this ratio with reference to a χ^2 distribution. Wald tests are more appropriate than *t*-tests for analyses that employ quasi-likelihood (QL) estimation instead of ordinary least squares (OLS) estimation, and can be used as an alternative to likelihood ratio tests (to compare the overall fit of different models) which are not available for non-linear models (e.g. Quinn & Keough, 2002; Rasbash et al., 2005).[56]

---

[54] For *p* predictor variables, the coefficients have the following nomenclature: $\beta_0, \beta_1, \beta_2, ...\beta_p$, where β_0 is the intercept, and $\beta_1, \beta_2, ...\beta_p$ are the partial regression coefficients of the added terms. β_0 is the log of the odds of a press on the lever associated with 2 pellets being made, relative to not (in the case of our dataset), when all the other predictors equal zero. $\beta_1, \beta_2, ...\beta_p$ are the log of the odds of the lever associated with 2 pellets being pressed for a single unit increase (i.e. whatever unit was used to measure that variable) in the associated predictor variable ($x_1, x_2, ...x_p$), when the other predictor variables are held constant (Quinn & Keough, 2002).

[55] Since it was a two-choice design with progression to the next trial conditional on a lever being pressed the response variable was binary, with a 'press' on one lever meaning 'no press' on the other, and *vice versa*.

[56] For examples of experimental papers which employ Wald tests in multilevel logistic regression modelling, see: Jones & Elgar (2004); Korsten et al. (2007); Magrath et al. (2003); Müller et al. (2003).

Multilevel analysis of the single-frequency probe session data: latency

When analysing *latency* as the response (y) variable, estimation was conducted using the default option of iterative generalised least squares (IGLS; no approximation procedures were necessary since the model is linear) (Rasbash et al., 2005). Likelihood ratio tests (LRT) were conducted to compare the fit of different models to the observed data. The LRT statistic is calculated from the following formula:

$$-2 \log L_1 - (-2 \log L_2)$$

Where 'L' refers to the 'likelihood', namely the probability of obtaining the observed data if the fitted model were true. The subscripts, i.e. '1' and '2', identify which of two different ('nested') models the likelihood statistic is derived from ('nested', in this sense, means one of the models is a restricted form of the other). The significance of the LRT statistic can then be obtained by reference to a χ^2 distribution, with the degrees of freedom equalling the number of 'extra' parameters which differentiate the two nested models[57]. The LRT therefore allows us to gauge whether the addition of particular parameters produces a model which better explains, or accounts for, the observed data: i.e. whether the modification has 'improved' the model. As well as LRT statistics, normal distribution tests were also occasionally used to assess the significance of single parameters (Rasbash et al., 2005).

Multilevel analysis of the single-frequency probe session data: model-building

Automated procedures of model selection (such as forwards, backwards or stepwise selection) aren't available in MLwiN, and in general such procedures are often criticised for producing models which can be misleading, and which can also differ depending on the specific automated selection procedure used (e.g. Grafen & Hails, 2002; Quinn & Keough, 2002). We employed a forwards, stepwise

[57] E.g. If the two models were the same, except Model '2' had the additional parameter of *treatment* , then there would be 1 d.f.

procedure, insofar as we built the model up from 'basics', but one which was manual, as opposed to automated, so that we were able to actively review the model at each stage, and make appropriate decisions. We also tended, where we could, towards parsimony, both in our selection of main effects, and when fitting higher-order interactions. This reduces the risk of finding significant effects by chance (especially when testing partial regression coefficients), but also reduces the risk of over-fitting a model: i.e. producing a model of such complexity that it is a very good fit to the data, but has little explanatory or predictive power (Quinn & Keough, 2002).

As Quinn & Keough (2002) note, when selecting predictors to add to a multiple regression model (be it multilevel, or otherwise) it's important to be aware of any multicollinearity – i.e. any correlations between the predictor variables – and if it does exist, to be sensitive to any impact it may have on the estimates of the coefficients, and hence hypothesis tests (Aiken & West, 1991; Kreft & De Leeuw, 1998). In a counterbalanced, orthogonal experimental design such as the one employed here, many of the main factors of interest will not be correlated, and so it is less of an obvious concern. However, if higher-order interactions are fitted, e.g. *'XZ'* in the equation:

$$Y = \beta_0 + \beta_1 X + \beta_2 Z + \beta_3 XZ$$

then these interactions may be correlated with their constituent conditional effects (i.e. in this example, *'XZ'* may be correlated with either *'X'* and/or *'Z'*) (Aiken & West, 1991). If multicollinearity in a model is moderate, then it is unlikely to be a problem (Zar, 1996), however if it is substantial, then it can have two important consequences: it can lead to instability in the model, with the addition/deletion of certain predictors leading to disproportionate changes in the coefficient value of any terms in the model with which they are correlated; and, in the case of interactions, it can lead to an over-inflation of the standard error of the constituent terms – e.g. in the example above, the addition of the interaction *'XZ'* could, if multicollinearity were present, result in an inflation of the standard error of the coefficient of *'X'* and/or the coefficient of *'Z'* (e.g. *'2.201(0.921)*X'* might change to *'2.201(1.251)*X'*, following the addition of *'XZ'*), risking a mistaken failure to reject

the null hypothesis (i.e. a Type II error) (e.g. Kreft & De Leeuw, 1998; Quinn & Keough, 2002).

The risk of multicollinearity between interactions and their constituent terms is substantially lowered if continuous predictor variables are centred (e.g. each datapoint is subtracted from the grand mean of that data series) (e.g. Aiken & West, 1991; Quinn & Keough, 2002), and, as mentioned earlier, we take this precaution in our analyses. Furthermore, following the addition of a new predictor term (e.g. the addition of *'W'*, in the above equation), we report any main effect of that term first, before building the model further by adding any interactions of interest featuring that term[58]. Finally, whenever we find effects of interest, but there is reason to suspect they may correlate with other terms in the equation, we test how robust such effects are by investigating them in more than one model (e.g. a simplified vs. more complex model).

Finally, a note on the interpretation of interaction terms, and their constituent conditional effects. When term *'X'* is added to a model, solely by itself (i.e. not a part of any interaction), it is generally termed a *main effect*, and its coefficient is an estimate of the constant, or average, effect of that variable (*'X'*) across the range of other variables in the model. If it is also added as part of an interaction (e.g. *'XZ'*), then what was the *main effect* (i.e. the term consisting only of *'X'*) now becomes a *conditional effect:* an estimate of the effect of *'X'* when *'Z'* equals zero (therefore, if *'Z'* is a centred continuous predictor, zero will be the mean of that predictor, whereas if *'Z'* is a categorical (dummy) variable, zero will be whichever categorical group has been assigned that value) (Aiken & West, 1991). By observing conditional effects, we can therefore interpret the meaning of interaction terms.

[58] N.B. If higher-order interactions (e.g. *'XZ'*) do substantially correlate with their lower-order constituent terms (e.g. *'X'*), this may change the coefficient estimate (and its error) of that constituent term (i.e. the coefficient of *'X'*), but will not effect the estimate of the higher-order interaction (i.e. the coefficient of *'XZ'*) (Aiken & West, 1991).

RESULTS

Training performance

All rats reached the initial criterion of three successive sessions in which the number of correct responses in free-choice trials was significantly greater than chance for both reference stimuli (i.e. 2kHz and 4kHz) in a binomial test. The mean number of training sessions to reach criterion was 26.6 (SEM: 1.4). There was no significant difference between prospective *treatment* groups in the number of training sessions taken to reach criterion (t_{14}=0.043, p=0.966), nor in the number of post-criterion training sessions prior to probe-testing (t_{14}=0.952, p=0.357)[59]. Figure 2.11 charts the percentage of 'correct' responses in the '1-pellet' and '2-pellet' reference trials, across training session. It indicates an initial bias towards pressing the '2-pellet' lever in both types of reference trial, although this attenuates with regard to the '1-pellet' reference tone as the sessions progress, and the *subjects* near criterion.

For reference, Figure 2.12 plots the mean accuracy in the reference trials in those training sessions conducted just prior to probe-testing in the two *measurement phases*. It indicates an improvement in accuracy across session in each of the two *measurement phases*, especially with regard to the '1-pellet' reference trials, and also indicates that on the day prior to probe testing, the mean accuracy across all the *subjects* was very similar in each of the two *measurement phases*.

[59] Nor in the total number of training sessions prior to probe-tesing (i.e. both pre- and post-criterion; p=0.726)

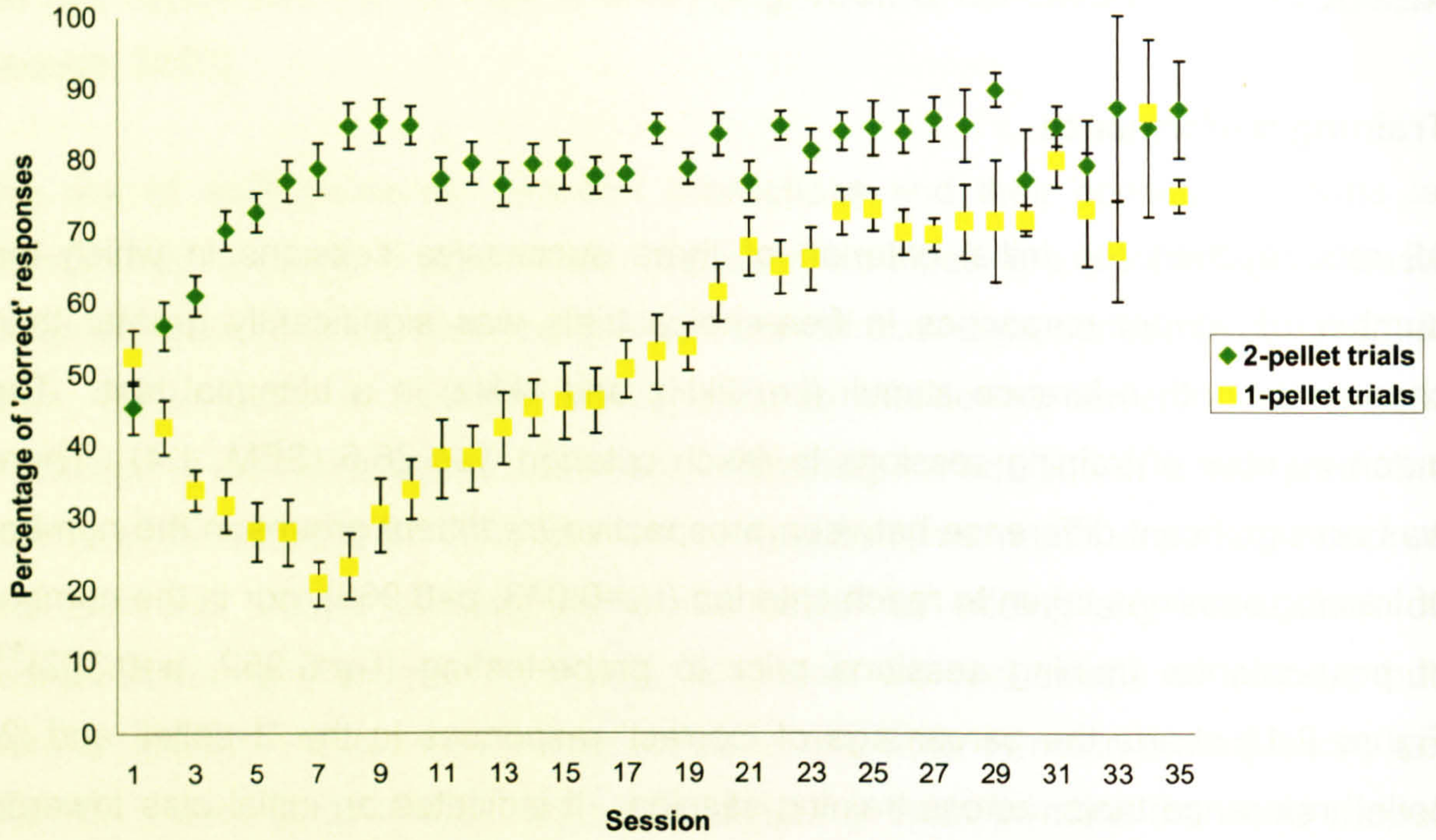

Figure 2.11 The mean percentage of 'correct' responses, across training session, for each of the two types of training trial, by quantity of associated food reinforcement (+/- 1SEM). The data is taken only from the free-choice trials, and only from *subjects* yet to complete criterion (i.e. the sample size reduces as session number progresses).

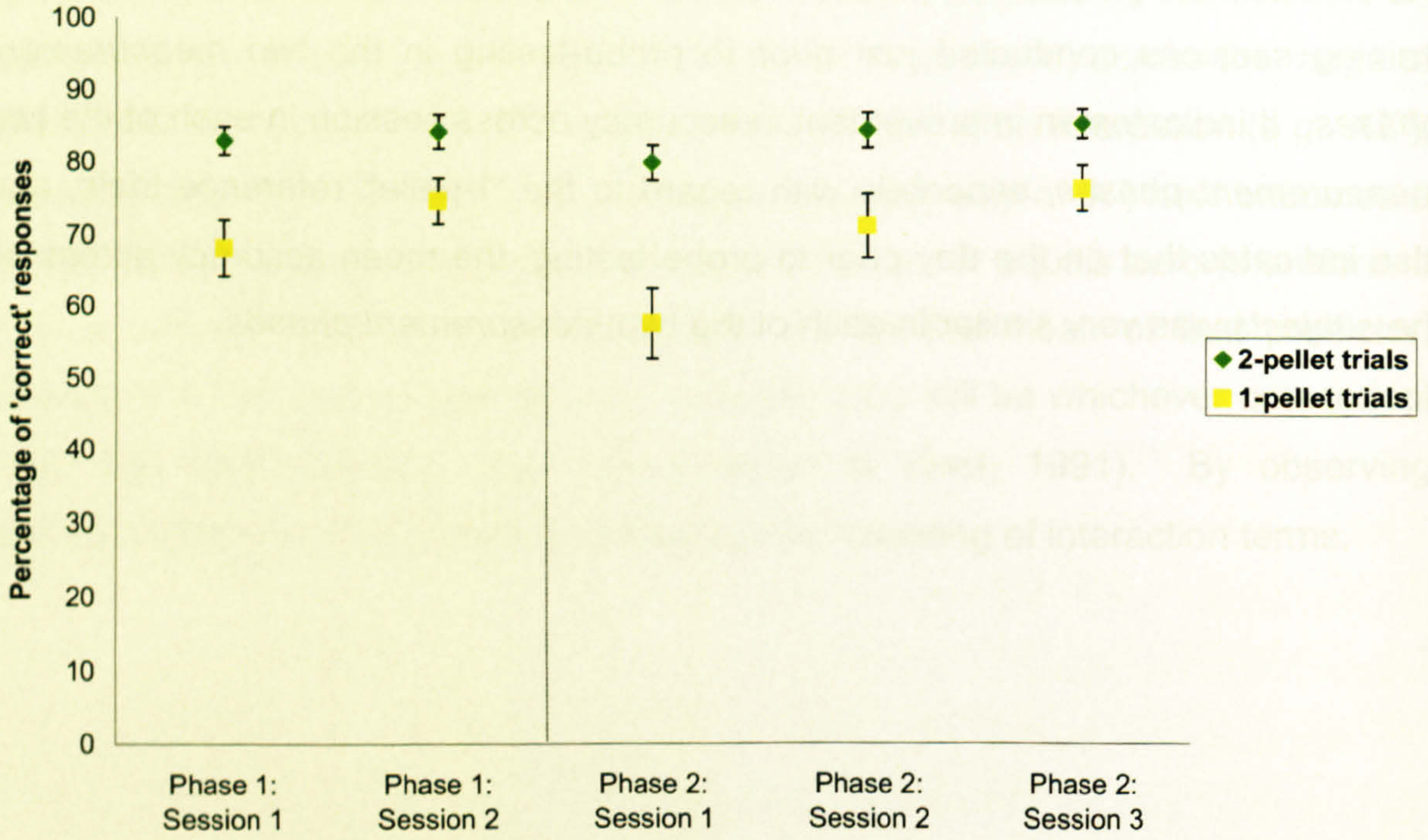

Figure 2.12 As Figure 2.11, but for each of the training sessions on the consecutive days immediately prior to single-frequency probe-testing in each of the two measurement phases.

Figure 2.13 plots the same data, but pooled across *measurement phase,* and summarised by *treatment* group. Analyses revealed that there were no significant differences between the *treatment* groups in the accuracy of responding to each type of reference tone, nor across all reference tones, either over the two training sessions prior to probe-testing in *phase 1* ('2-pellet' trials: t_{14}=0.311, p=0.760; '1-pellet' trials: t_{14}=0.289, p=0.777; all trials: t_{14}=0.402, p=0.694), nor over the three training sessions prior to probe-testing in *phase 2* ('2-pellet' trials (cube-transformed): t_{14}=0.686, p=0.504; '1-pellet' trials: t_{14}=0.104, p=0.919; all trials: t_{14}=0.550, p=0.591). Furthermore, whilst Figure 2.13 suggests that the *UHT* group were slightly more accurate than the *Control* group in the second *measurement phase,* whereas the reverse was true in the first *measurement phase,* some exploratory analyses revealed that this interaction was not signficant (repeated-measures GLMs, with *measurement phase* as the within-subjects factor and *treatment* as the between-subjects factor, were conducted separately for '2-pellet' trials, '1-pellet' trials, and for all trials: these found p>0.05 for all main effects and interactions).

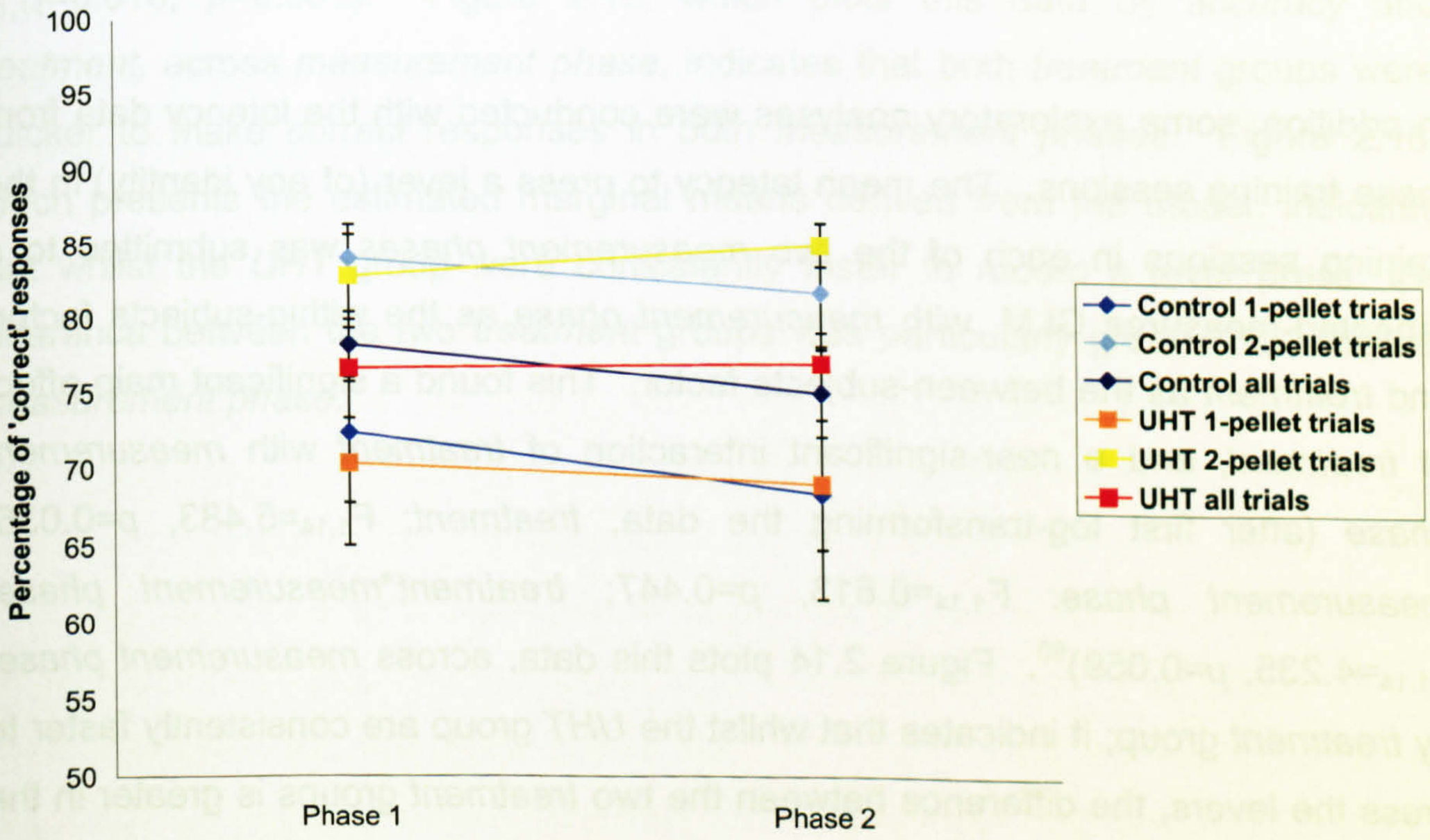

Figure 2.13 The mean percentage of 'correct' responses in the training sessions in each of the two *measurement phases*, by *treatment*. The data is summarised by trial type (+/- 1SEM).

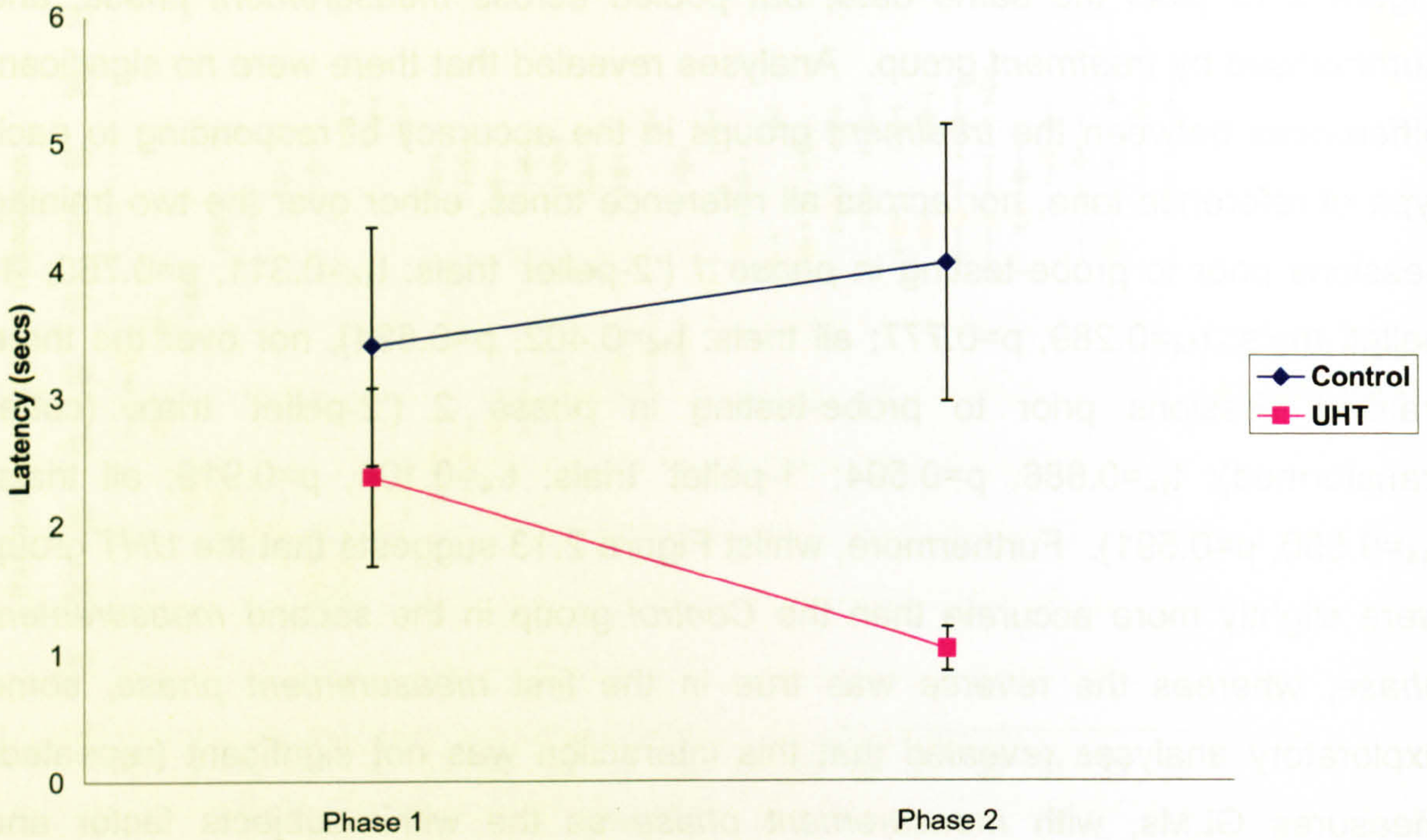

Figure 2.14 The mean latency to record *any* lever press response in the training sessions which took place on the consecutive days immediately prior to single-frequency probe testing, summarised by *measurement phase* (two such training sessions took place in *phase 1*, and three took place in *phase 2*), by *treatment* group (+/- 1SEM).

In addition, some exploratory analyses were conducted with the latency data from these training sessions. The mean latency to press a lever (of any identity) in the training sessions in each of the two *measurement phases* was submitted to a repeated-measures GLM, with *measurement phase* as the within-subjects factor, and *treatment* as the between-subjects factor. This found a significant main effect of *treatment,* and a near-significant interaction of *treatment* with *measurement phase* (after first log-transforming the data, *treatment:* $F_{1,14}$=5.483, p=0.035; *measurement phase:* $F_{1,14}$=0.613, p=0.447; *treatment*measurement phase:* $F_{1,14}$=4.235, p=0.059)[60]. Figure 2.14 plots this data, across *measurement phase,* by *treatment* group; it indicates that whilst the *UHT* group are consistently faster to press the levers, the difference between the two *treatment* groups is greater in the

[60] N.B. whilst our inspection of the plotted residuals suggested the variance was reasonably homogenous, the residuals from *Phase 2* failed a formal test of homogeneity (Levene's: p=0.021); this could not be remedied by any of the transformations we explored, at least not without compromising other assumptions of the model.

second *measurement phase,* with the *UHT* group considerably faster, and the *Control* group slower, in *phase 2* compared to *phase 1.*

The latency to record a lever press response in only those trials associated with 2 pellets of food (i.e. the trials in which 2kHz was presented for the *2kHz=2pell contingency* group, and the trials in which 4kHz was presented for the *4kHz=2pell contingency* group) in the training sessions prior to single-frequency probe-testing was also submitted to a repeated-measures GLM, with both *measurement phase* and *accuracy of response* (i.e. *'correct'* or *'incorrect'*) as within-subject factors, and *treatment* as a between-subjects factor. This found a significant main effect of *accuracy of response*, and a significant interaction of *measurement phase* with *treatment*; all other main effects and interactions were non-significant, although *treatment* neared significance at the 0.05 level (after first log-transforming the data, *treatment:* $F_{1,14}$=3.712, p=0.075; *measurement phase:* $F_{1,14}$=0.806, p=0.384; *accuracy:* $F_{1,14}$=9.110, p=0.009; *measurement phase*treatment:* $F_{1,14}$=6.456, p=0.024; *accuracy*treatment:* $F_{1,14}$=1.256, p=0.281; *measurement phase*accuracy*: $F_{1,14}$=1.562, p=0.232; *measurement phase*accuracy*treatment:* $F_{1,14}$=0.016, p=0.901). Figure 2.15, which plots this data by *accuracy* and *treatment,* across *measurement phase*, indicates that both *treatment* groups were quicker to make *correct* responses in both *measurement phases.* Figure 2.16, which presents the estimated marginal means derived from the model, indicates that whilst the *UHT* group were consistently faster to record a lever press, the difference between the two *treatment* groups was particularly great in the second *measurement phase.*

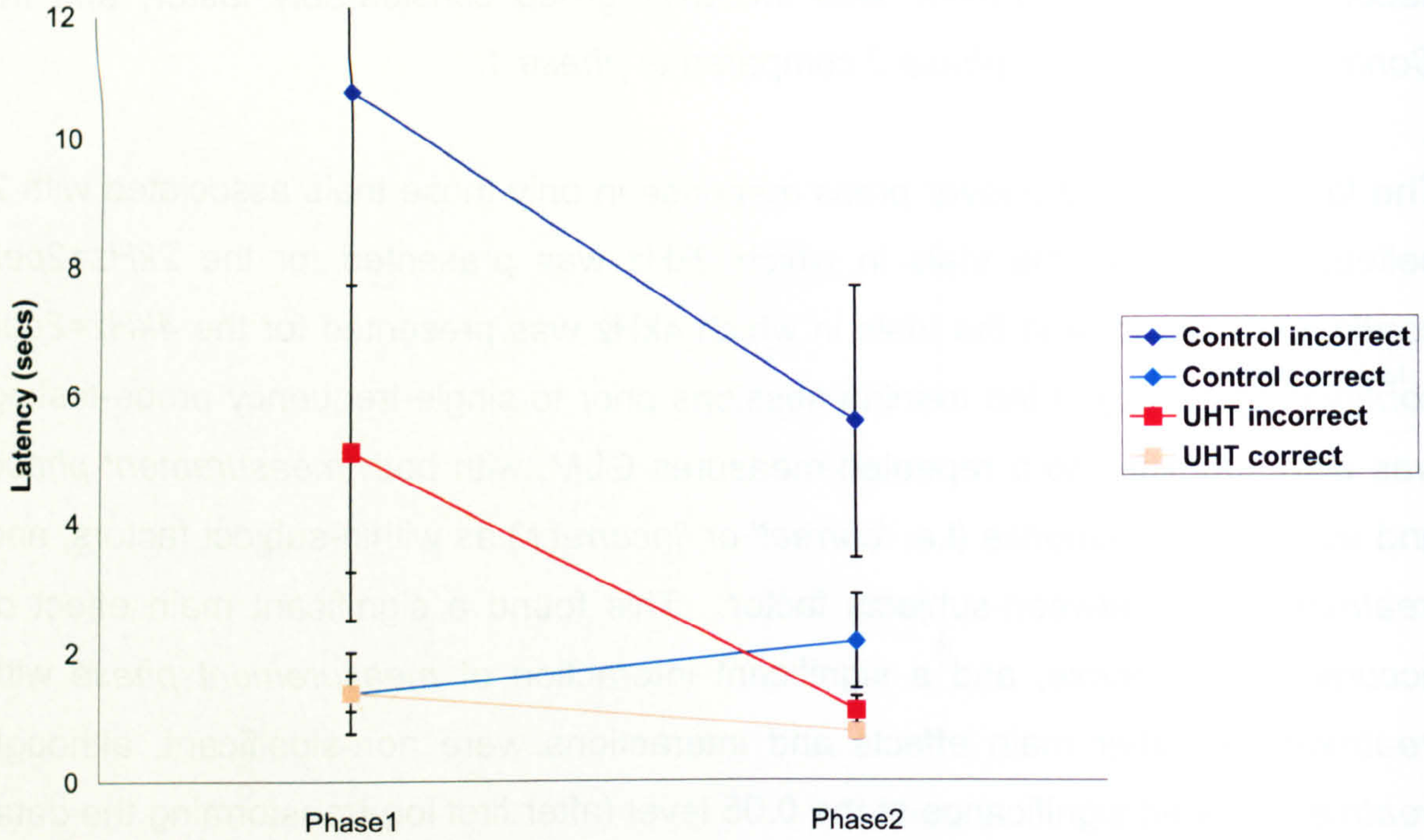

Figure 2.15 '2-pellet' trials only. The mean latency to record a 'correct' (i.e. press the '2-pellet' lever) or 'incorrect' (i.e. press the '1-pellet' lever) response in the trials associated with 2 pellets of food in the training sessions which took place on the consecutive days immediately prior to single-frequency probe testing, summarised by *measurement phase* (two such training sessions took place in *phase 1*, and three took place in *phase 2*), by *treatment* group (+/- 1SEM).

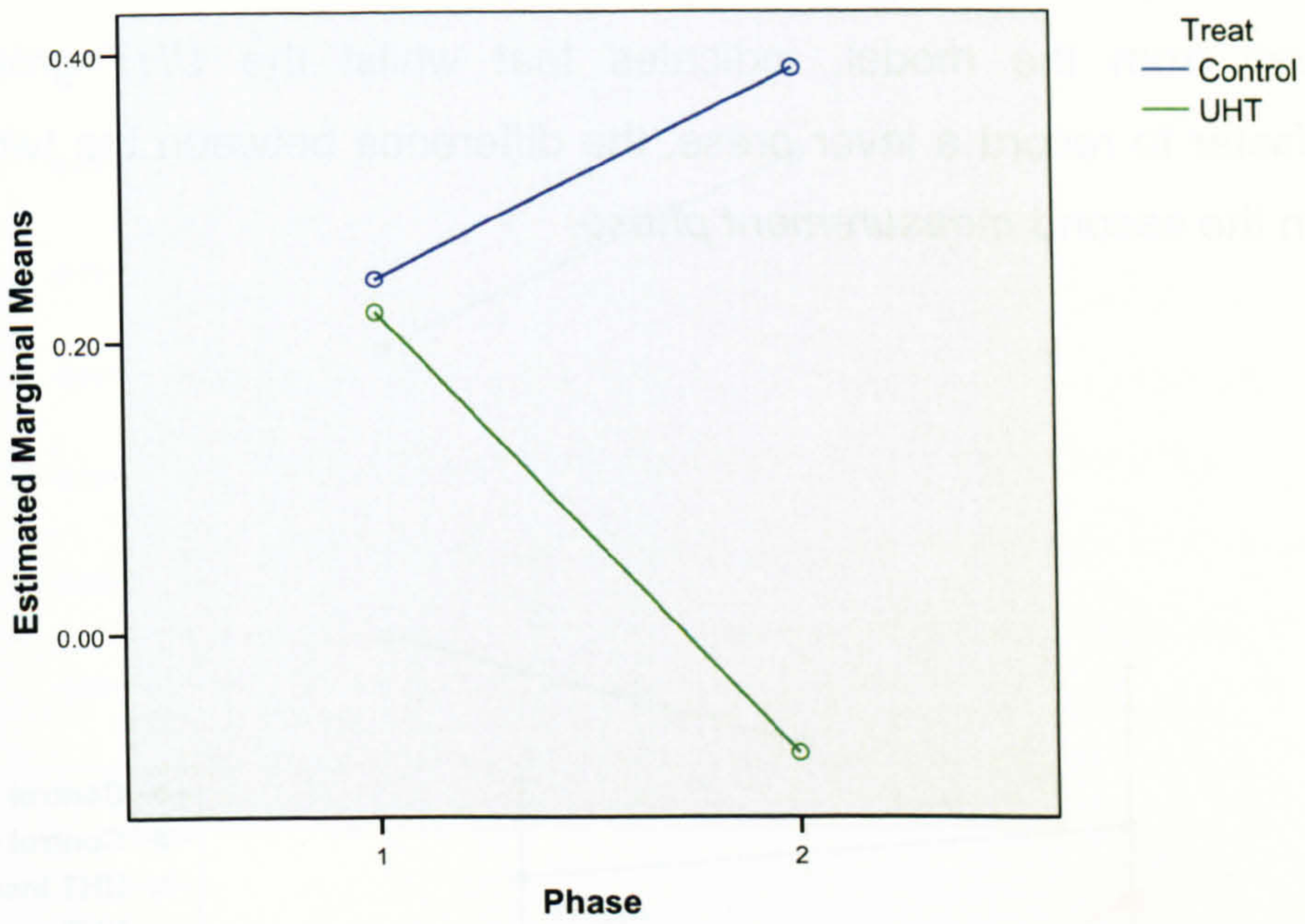

Figure 2.16 The estimated marginal means of the (log-transformed) mean latency to record a lever press response in the trials associated with 2 pellets of food, across *measurement phase* for each *treatment* group, derived from the repeated-measures GLM discussed in the text.

Finally, the latency to record a lever press response in only those trials associated with 1 pellet of food (i.e. the trials in which 4kHz was presented for the *2kHz=2pell contingency* group, and the trials in which 2kHz was presented for the *4kHz=2pell contingency* group) in the training sessions prior to single-frequency probe-testing was also submitted to a repeated-measures GLM, of the same design as that employed above. The analysis found significant main effects of *accuracy,* and of *treatment*; whilst all other main effects and interactions were non-significant (after first negative inverse-transforming the data, *treatment:* $F_{1,14}$=4.747, p=0.047; *measurement phase:* $F_{1,14}$=0.130, p=0.724; *accuracy:* $F_{1,14}$=15.145, p=0.002; *measurement phase*treatment:* $F_{1,14}$=1.723, p=0.210; *accuracy*treatment:* $F_{1,14}$=0.212, p=0.652; *measurement phase*accuracy*: $F_{1,14}$=2.216, p=0.159; *measurement phase*accuracy*treatment:* $F_{1,14}$=0.107, p=0.749). Figure 2.17,

which plots this data by *accuracy* and *treatment,* across *measurement phase*, indicates that both *treatment* groups were quicker to make *correct* responses in both *measurement phases*. Figure 2.18, which presents the estimated marginal means derived from the model, indicates that whilst the *UHT* group were consistently faster to record a lever press, the difference between the two groups was greater in the second *measurement phase.*

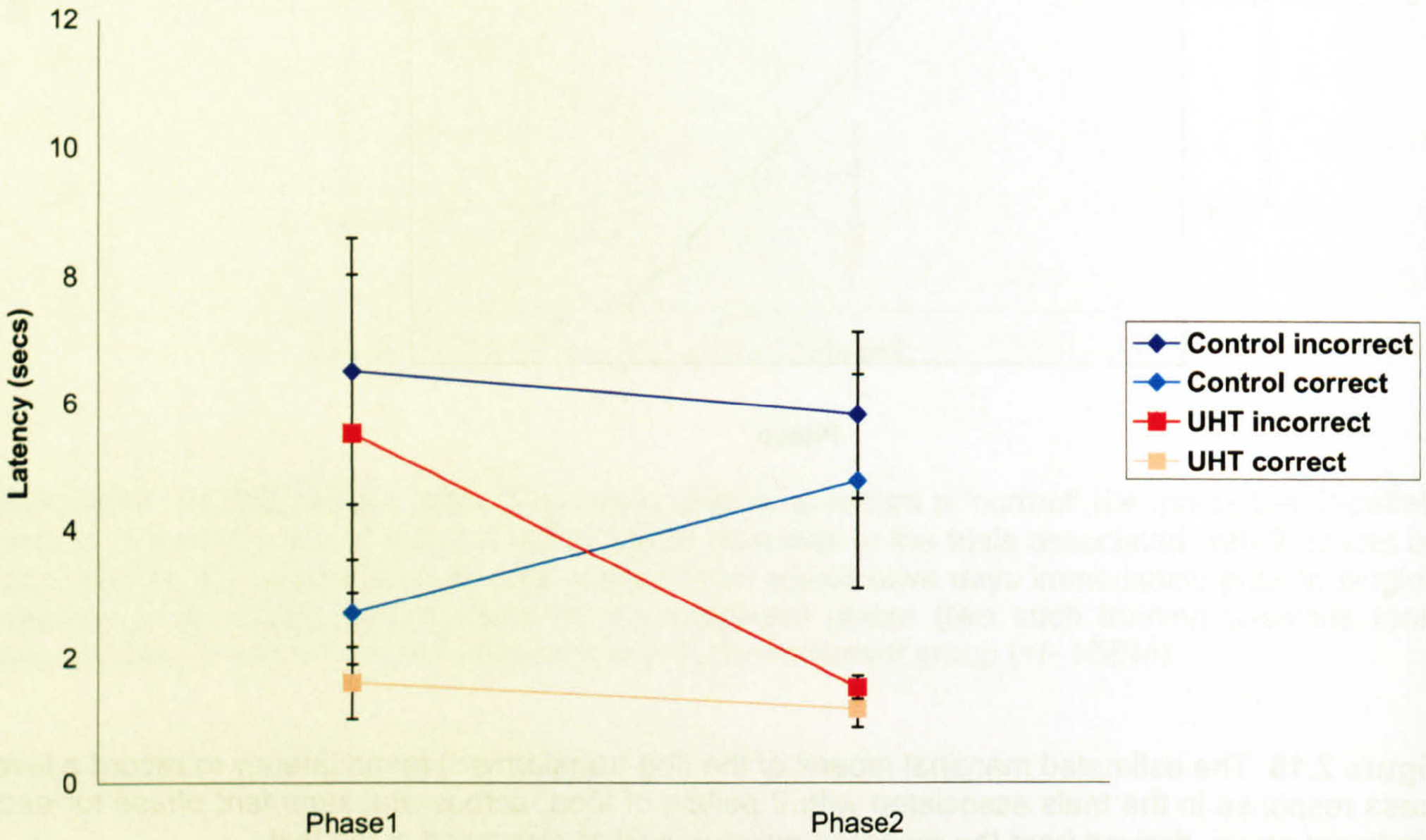

Figure 2.17 '1-pellet' trials only. As Figure 2.15, but for the trials associated with 1 pellet of food only.

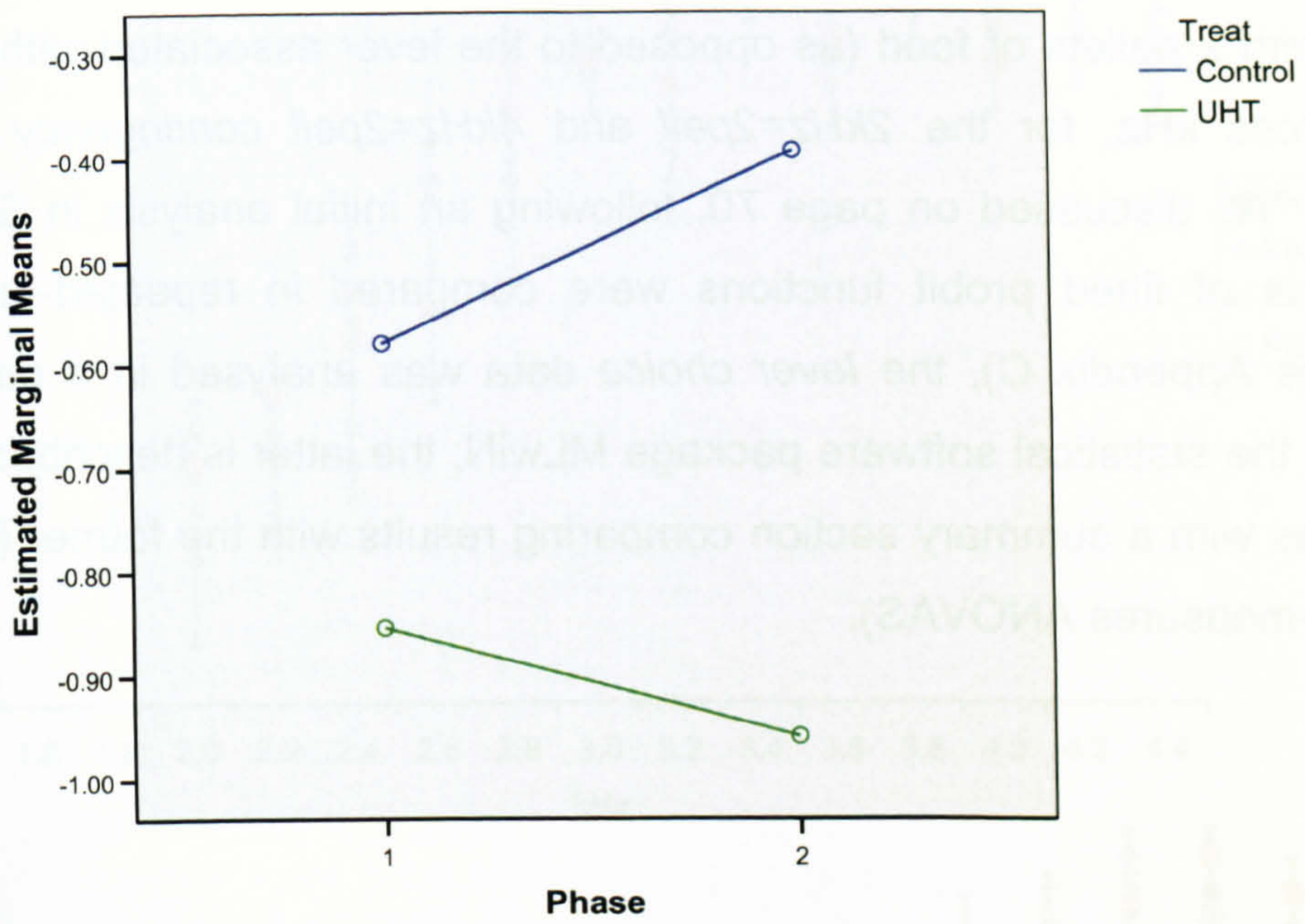

Figure 2.18 The estimated marginal means of the (negative-inverse-transformed) mean latency to record a lever press response in the trials associated with 1 pellet of food, across *measurement phase* for each *treatment* group, derived from the repeated-measures GLM discussed in the text.

Summary of the analyses of training performance

All rats reached the initial criterion, and the two *treatment* groups did not differ in the accuracy of their responding in the training sessions administered just prior to probe-testing. Whilst the *UHT* group were generally faster to record a lever press response in these training sessions (i.e. those administered in each *measurement phase*), the difference in latency between the two *treatment* groups was generally greater in the second *measurement phase* (to a significant level with regard to the '2-pellet' trials, and to a near-significant level across all reference trials). A variety of analyses confirmed that the decrease in latency (i.e. increase in response speed) across *measurement phase* for the *UHT* group did not occur at the expense of accuracy.

Single-frequency test sessions: lever choice

Figure 2.19 and Figure 2.20 chart the mean percentage of presses on the lever associated with 2 pellets of food (as opposed to the lever associated with 1 pellet of food) across kHz, for the *2kHz=2pell* and *4kHz=2pell contingency* groups, respectively. As discussed on page 70, following an initial analysis in SPSS (in which aspects of fitted probit functions were compared in repeated-measures ANOVAs; see Appendix C), the *lever choice* data was analysed in a multi-level model, using the statistical software package MLwiN; the latter is described below, and concludes with a summary section comparing results with the former (i.e. with the repeated-measures ANOVAS).

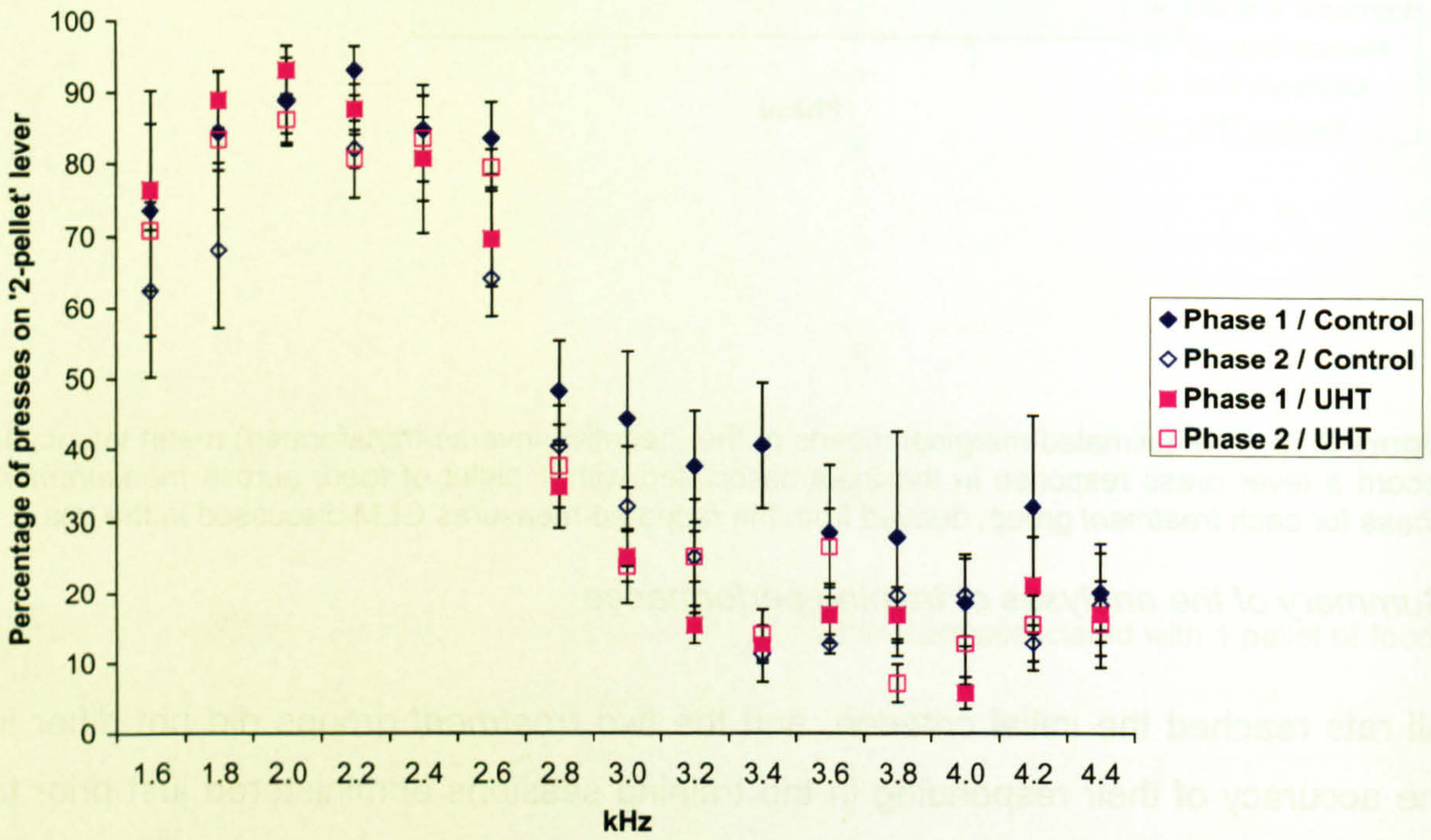

Figure 2.19 ***2kHz=2pell.*** The mean percentage of lever presses made on the '2-pellet' lever (as opposed to the '1-pellet' lever) for the *2kHz=2pell contingency* group, by *measurement phase / treatment* group, across kHz (+/- 1 SEM. N.B. the means and SEM are derived from data summarised at the *subject*-level; in addition, the data pertaining to the 'reference tones' (i.e. 2kHz and 4kHz) are taken from the non-reinforced trials only).

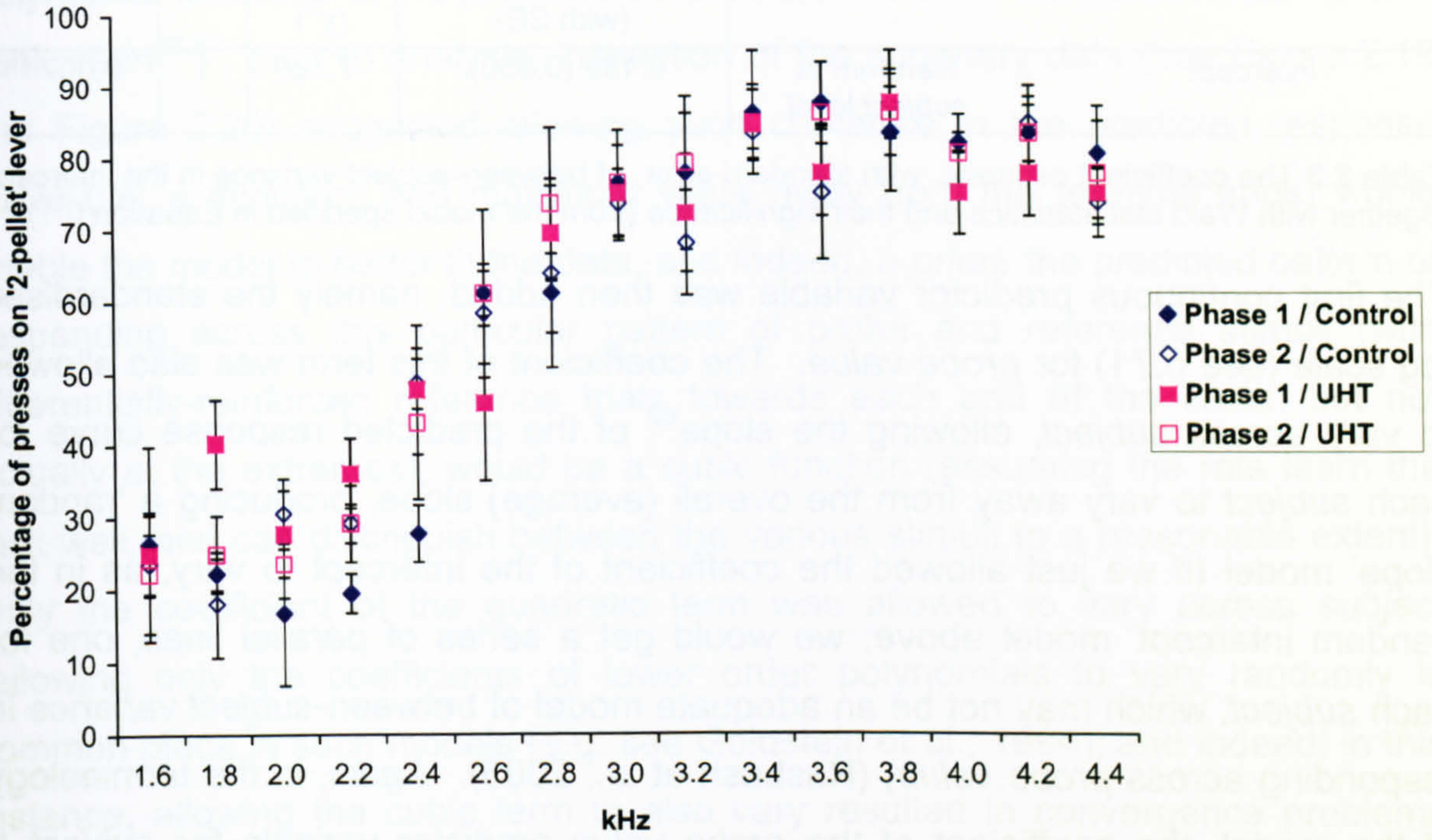

Figure 2.20 ***4kHz=2pell.*** As Figure 2.19, but for the *4kHz=2pell contingency* group only.

Multi-level multiple logistic regression in MLwiN

As discussed on p.74, the *lever choice* data in the single-frequency probe sessions were analysed in a multi-level model using MLwiN, with *trial* (n=8,626) nested within *subject* (n=16).

Firstly, a simple 'random intercept' model was fitted, in which the intercept for each *subject* was allowed to vary from the overall (average) intercept (β_0). In the terminology of the model, the intercept for *subject 'j'* differs from the overall (fixed) intercept by u_{0j}, which is a random quantity assumed to follow a normal distribution (Rasbash et al., 2005). As shown in Table 2.3, the estimated coefficients (of σ^2_u in this instance) from this very simple model reveal that the probability of pressing the lever associated with 2 pellets of food varies significantly across *subject*. Equation 0.1, in the Appendix, shows the relevant equations with estimated coefficients[61].

[61] Whilst building our model, we will occasionally refer to selected equations we have copied into the Appendix. It's not necessary to refer to all (or any) of these: they are simply provided for reference, should the reader wish to study other aspects of the model not reported in the main text.

Parameter		Coefficient estimate (with SE)	Wald (χ^2)	Df	*P*
Intercept	Random at *subject* level	0.135 (0.050)	7.169	1	0.007 **

Table 2.3 The coefficient estimate, with standard error, of between-*subject* variance in the intercept, together with Wald test statistics and their significance (from the model specified in Equation 0.1).

The first continuous predictor variable was then added, namely the standardised log scale (see p.71) for *probe value*. The coefficient of this term was also allowed to vary across *subject,* allowing the slope[62] of the predicted response curve for each *subject* to vary away from the overall (average) slope, producing a 'random slope' model (if we just allowed the coefficient of the intercept to vary, as in the 'random intercept' model above, we would get a series of parallel lines, one for each *subject*, which may not be an adequate model of between-*subject* variance in responding across *probe value*) (Rasbash et al., 2005). Again, in the terminology of the model, the coefficient of the *probe value* predictor variable for *subject 'j'* varies from the overall (fixed) coefficient by u_{1j}, a random quantity assumed to follow a normal distribution. Table 2.4 shows that there is a highly-significant fixed (average, or overall) effect of *probe value* on lever choice, with a greater probability of pressing the lever associated with 2 pellets of food as *probe value* increases (i.e. as the scale moves towards the stimulus associated with 2 pellets of food; more specifically, a one unit increase in *probe value* increases the log odds of a press being made on the lever associated with 2 pellets of food by 2.766); there is also a significant amount of between-*subject* variation in lever choice across *probe value*.

Parameter		Coefficient estimate (with SE)	Wald (x2)	Df	P
Intercept	Random at *subject* level	0.164 (0.062)	7.065	1	0.008 **
Probe value	Random at *subject* level	0.524 (0.211)	6.185	1	0.013 *
	Fixed	2.766 (0.193)	205.169	1	<0.001 **

Table 2.4 The coefficient estimates, with standard error, of fixed and random parts of a simple 'random slope' model, together with Wald test statistics and their significance (as specified in Equation 0.2, in the Appendix).

[62] Note, some of the terminology we use to describe our model (e.g. 'slope') is perhaps more commonly associated with linear models (such as the analysis in which we model *latency* as the response variable, below), rather than non-linear models (e.g. modelling a binary distribution). However, this is in keeping with others' description of non-linear models (e.g. Rasbash et al., 2005), has an intuitive appeal, and introduces a language we can also use to describe the later linear model.

Polynomial functions of the *probe value* predictor variable were then fitted up to a cubic term[63,64]; prior to analysis, inspection of the summary data (see Figure 2.19 and Figure 2.20) suggested allowing such curvature in the predicted response curve (i.e. a minimum and maximum value away from the terminal ends) would enable the model to better fit the data, and indeed, *a priori,* the predicted pattern of responding across this particular pattern of probe and reference stimuli (with differentially-reinforced reference trials towards each end of the scale, but not actually at the extremes), would be a cubic function (assuming the rats learn the task well, and can distinguish between the various stimuli to a reasonable extent). Only the coefficient of the quadratic term was allowed to vary across *subject* (allowing only the coefficients of lower order polynomials to vary randomly is common-place in such models (e.g. see Goldstein et al., 1994), and indeed, in this instance, allowing the cubic term to also vary resulted in convergence problems when further terms were later added to the model (e.g. Nezlek, 2001)). As Table 2.5 indicates, the *probe value* terms (i.e. up to a cubic power) accounted for a significant amount of the overall (fixed) variance in lever choice: i.e. across all *subjects*, and all *trials,* expressing *probe value* up to a cubic power provides a very good fit to the data. In addition, those *probe value* terms allowed to vary across *subject* indicated a significant amount of between-*subject* variance in the probability of *lever choice* across *probe value*. Figure 2.21 and Figure 2.22 plot the predicted probability of pressing the '2-pellet' lever for each subject, with *probe value* and *kHz* on the x-axes, respectively.

[63] N.B. when fitting a cubic polynomial regression model, the lower order (i.e. linear and quadratic) terms must be included to respect marginality (e.g. Grafen & Hails, 2002).

[64] Note that the linear *probe value* term here was centred, as were all the other continuous predictor variables; however, whilst the quadratic and cubic powers were derived from this centred term, they were not themselves centred.

Parameter		Coefficient estimate (with SE)	Wald ($\chi 2$)	Df	*P*
Intercept	Random at *subject* level	0.242 (0.093)	6.770	1	0.009 **
Probe value	Random at *subject* level	0.522 (0.212)	6.048	1	0.014 *
	Fixed	4.858 (0.233)	432.896	1	<0.001 **
$(Probe\ value)^2$	Random at *subject* level	2.558 (1.024)	6.327	1	0.012 *
	Fixed	-0.834 (0.425)	3.847	1	0.049 *
$(Probe\ value)^3$	Fixed	-5.993 (0.368)	265.551	1	<0.001 **

Table 2.5 The coefficient estimates, with standard error, of fixed and random parts of a 'random slope' model with polynomial terms, together with Wald test statistics and their significance (as specified in Equation 0.3, in the Appendix).

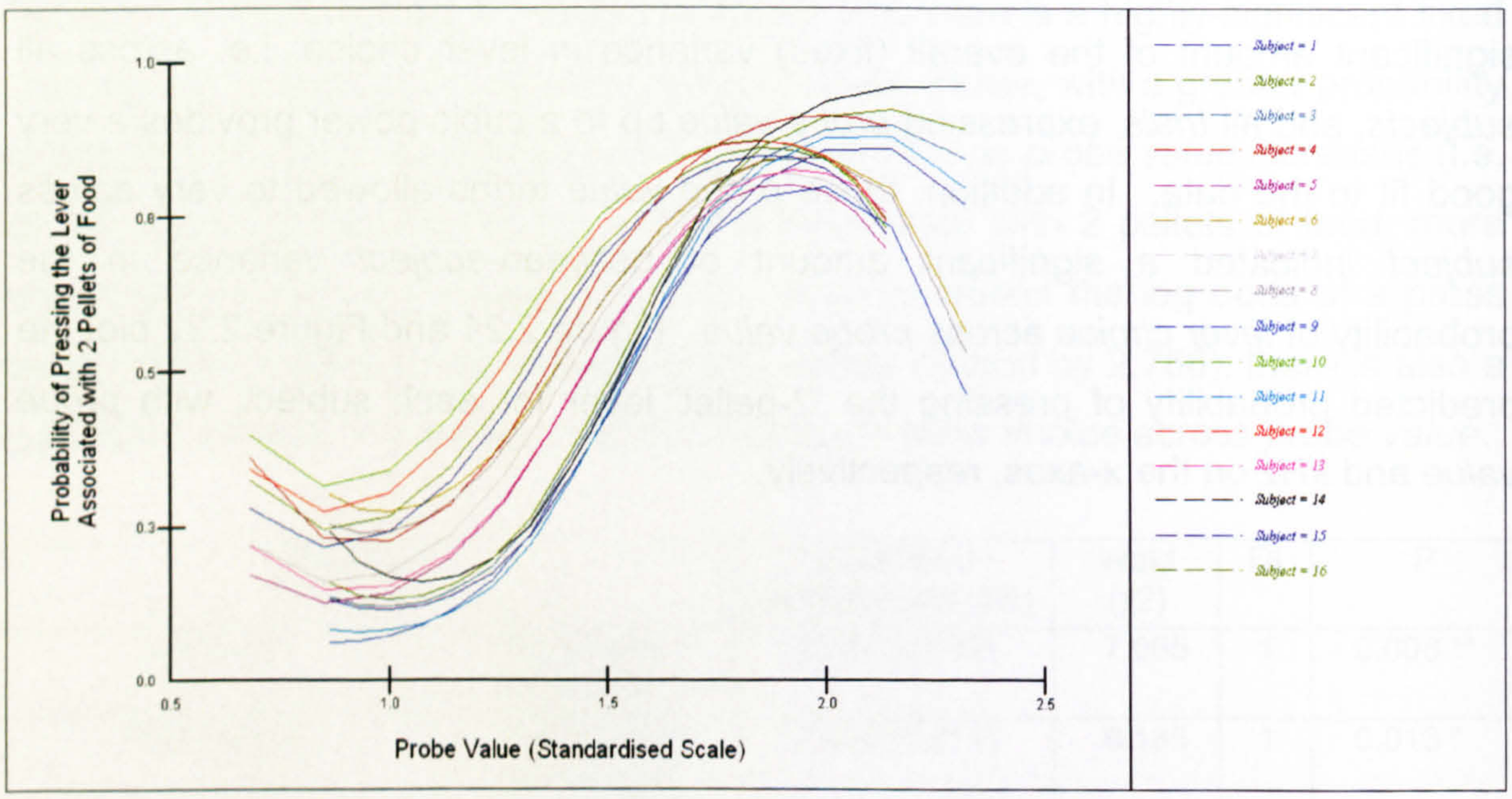

Figure 2.21 Individual rats' probability of pressing the lever associated with 2 pellets of food (as opposed to the lever associated with 1 pellet of food) across different values of the probe stimuli (a standardised scale, with *1.0* corresponding to the reference tone associated with 1 pellet of food, and *2.0* corresponding to the reference tone associated with 2 pellets of food). These predicted lines were generated from the model and estimates in Equation 0.3, and include data from *all* single-frequency probe sessions (i.e. from both *measurement phases*). Note, both this chart, and Figure 2.22, are included to illustrate the general shape of the predicted response curves: it is not necessary to identify individual *subjects,* and the following guide is simply provided for optional reference*: control / 2kHz=2pell → subjects 3, 6, 8 & 11; control / 4kHz=2pell → subjects 4, 5, 7 & 12; UHT / 2kHz=2pell → subjects 1, 9, 14 & 16; UHT / 4kHz=2pell → subjects 2, 10, 13 & 15.*

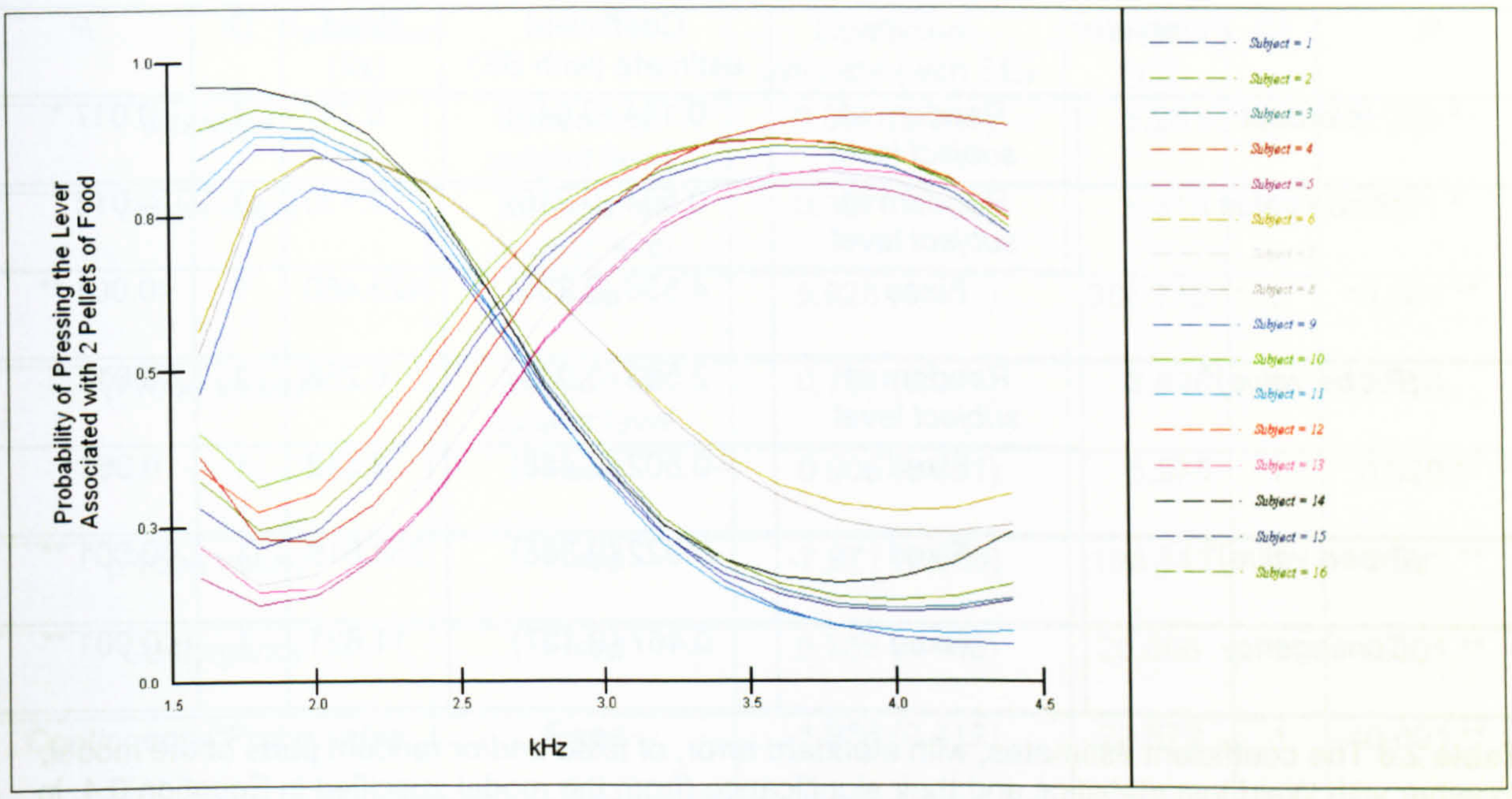

Figure 2.22 As Figure 2.21, but with kHz on the x-axis.

The categorical predictor variables of main interest (*contingency, measurement phase* and *treatment*), were then systematically added to the model as fixed effects.

Contingency was first added to the model, initially as a main effect, with the *2kHz=2pell* group assigned the reference category, with a value of *'0'*, and the *4kHz=2pell* group assigned a value of *'1'*. As Table 2.6 shows, the *contingency* group *4kHz=2pell* had a significantly higher overall probability of pressing the lever associated with 2 pellets[65]. In addition, there is a slight decrease in the amount of unexplained *subject*-level variation in the intercept (i.e. *contingency* has accounted for some of this variance).

[65] Since the coefficient was positive, and this group was assigned the value with the higher numerical value.

Parameter		Coefficient estimate (with SE)	Wald (χ2)	Df	*P*
Intercept	Random at *subject* level	0.114 (0.048)	5.747	1	0.017 *
Probe value	Random at *subject* level	0.534 (0.216)	6.130	1	0.013 *
	Fixed	4.836 (0.235)	423.460	1	<0.001 **
(Probe value)^2	Random at *subject* level	2.583 (1.033)	6.256	1	0.012 *
	Fixed	-0.802 (0.428)	3.512	1	0.061
(Probe value)^3	Fixed	-5.922 (0.368)	258.819	1	<0.001 **
Contingency	Fixed	0.467 (0.137)	11.621	1	<0.001 **

Table 2.6 The coefficient estimates, with standard error, of fixed and/or random parts of the model, together with Wald test statistics and their significance (from the model specified in Equation 0.4, in the Appendix).

Contingency, specified solely as a main effect, can only modify the intercept, but it would also be useful to allow it to modify the pattern of *lever choice* across *probe value* (i.e. the 'slope' of the predicted response curve). To allow this, we introduced two-way interaction terms featuring *contingency* with each of the *probe value* predictors. As Table 2.7 shows, these interactions were highly significant, indicating that the pattern of responding across *probe value* does indeed differ between the two *contingency* groups (see later plots); in addition, a little more of the previously unexplained *subject*-level variation has been accounted for.

Parameter		Coefficient estimate (with SE)	Wald (x2)	Df	*P*
Intercept	Random at *subject* level	0.091 (0.040)	5.281	1	0.022 *
Probe value	Random at *subject* level	0.307 (0.133)	5.313	1	0.021 *
	Fixed	5.928 (0.301)	386.712	1	<0.001 **
(Probe value)^2	Random at *subject* level	0.731 (0.374)	3.826	1	0.050
	Fixed	0.906 (0.391)	5.375	1	0.020 *
(Probe value)^3	Fixed	-7.971 (0.565)	198.841	1	<0.001 **
Contingency	Fixed	0.765 (0.168)	20.666	1	<0.001 **
*Contingency*Probe value*	Fixed	-1.956 (0.412)	22.572	1	<0.001 **
Contingency(Probe value)*^{2}	Fixed	-2.680 (0.549)	23.863	1	<0.001 **
Contingency(Probe value)*^{3}	Fixed	3.204 (0.768)	17.407	1	<0.001 **

Table 2.7 As Table 2.6, but derived from a modified model, as specified in Equation 0.5, in the Appendix.

Measurement phase was then added to the model, initially as a main effect, with *Phase 1* (i.e. the measurement phase at baseline) the reference category, assigned a value of *'0'*, and *Phase 2* assigned a value of *'1'*. As Table 2.8[66] indicates, there was a significantly lower overall probability of pressing the lever associated with 2 pellets of food in the second *measurement phase.*

Parameter		Coefficient estimate (with SE)	Wald (x2)	Df	*P*
Measurement phase	Fixed	-0.142 (0.052)	7.302	1	0.007 **

Table 2.8 As Table 2.6, but derived from a modified model, as specified in Equation 0.6, in the Appendix.

[66] Note, that whilst Table 2.8 only presents details relating to one term, the model from which it was derived still contains all the terms introduced earlier in the model-building process (i.e. those listed in Table 2.7; see Equation 0.6, in the Appendix).

Interactions of *measurement phase* with both *probe value* and *contingency* were then explored in the model (see Table 2.9 and Table 2.10)[67]. The interactions of *measurement phase* with *probe value* were not significant (in a variety of models we investigated), indicating no change in the shape of the response curve across *probe value* between the two *measurement phases*, whilst the interaction between *measurement phase* and *contingency* was significant; since the three-way interaction between all three of these terms was not significant only *measurement phase*contingency* was left in the model. As Table 2.10 shows, further investigation of the conditional effect of *measurement phase*, by changing which of the two *contingency* groups *(2kHz=2pell* or *4kHz=2pell*) was assigned a value of zero in the two-way interaction of *measurement phase* with *contingency* (as discussed in the Method: Data Analysis section), indicated that the *2kHz=2pell contingency* group had a considerably lower probability of pressing the '2-pellet' lever, and the *4kHz=2pell contingency* group a slightly higher probability of pressing the '2-pellet' lever, in the second *measurement phase*, compared to the first[68].

The final term of main interest, *treatment*, was then added to the model, initially as a main effect; the *control* group was assigned the reference category, with a value of *'0'*, whilst the *UHT (*i.e. *unpredictably-housed)* group was assigned a value of *'1'*. As Table 2.11 indicates, the main effect of *treatment* on the probability of *lever choice* was not significant.

---

[67] NB these tables have rows separated by spaces to signify that not all are from the same model: I.e. they are derived from models which differ slightly from each other in their specification, as detailed in the Table's legend.

[68] Alternatively, it's possible to calculate this from the coefficient estimates: I.e. the coefficient of the conditional effect of *measurement phase* (N.B. *phase 1* is the reference category), when the *contingency* group *2kHz=2pell* is assigned a value of zero in the two-way interaction *measurement phase*contingency*, is -0.351: I.e. the *2kHz=2pell* group have a lower probability (we can tell this from the minus sign) of pressing the '2-pellet' lever in the second *measurement phase*, compared to the first. If we add the coefficient of the interaction term (0.395) to this estimate, we get 0.044: the effect of *measurement phase* on the *4kHz=2pell contingency* group.

Parameter		Coefficient estimate (with SE)	Wald (χ2)	Df	*P*
Measurement phase probe value*	Fixed	0.105 (0.267)	0.154	1	0.695
Measurement phase (probe value)^2*	Fixed	-0.508 (0.271)	3.499	1	0.061
Measurement phase (probe value)^3*	Fixed	-0.314 (0.611)	0.264	1	0.607
Measurement phase Contingency*	Fixed	0.395 (0.105)	14.091	1	<0.001 **
Measurement phase contingency*probe value*	Fixed	-0.305 (0.608)	0.253	1	0.615
Measurement phase contingency*(probe value)^2*	Fixed	0.121 (0.682)	0.032	1	0.859
Measurement phase contingency*(probe value)^3*	Fixed	1.160 (1.536)	0.571	1	0.450

Table 2.9 As Table 2.6, but derived from a modified model, as specified in Equation 0.6 (with the addition of all terms necessary to respect marginality).

Parameter		Coefficient estimate (with SE)	Wald (χ2)	Df	*P*
Measurement phase	Fixed	-0.351 (0.077)	20.903	1	<0.001 **
Measurement phase Contingency (2kHz=2pell = ref. category)*	Fixed	0.395 (0.105)	14.091	1	<0.001 **
Measurement phase	Fixed	0.044 (0.072)	0.371	1	0.542
Measurement phase Contingency (4kHz=2pell = ref. category)*	Fixed	-0.395 (0.105)	14.091	1	<0.001 **

Table 2.10 As Table 2.6, but derived from a modified model, as specified in Equation 0.7.

Parameter		Coefficient estimate (with SE)	Wald (χ2)	Df	*P*
Treatment	Fixed	-0.062 (0.133)	0.215	1	0.643

Table 2.11 As Table 2.6, but derived from a modifed model, as specified in Equation 0.8, in the Appendix.

The effect of *treatment* was then explored in a variety of interactions (tested in a number of alternative models): see Table 2.12 and Table 2.13. The interaction of *treatment* with *measurement phase* was highly significant, and further investigation of the conditional effect of *treatment,* and an examination of the relevant coefficients, revealed that the *UHT* group had a lower probability of pressing the '2-pellet' than the *control* group in the first *measurement phase*, but a higher probability of pressing this lever than the *control* group in the second *measurement phase* (see Table 2.13). In addition, the three-way interaction between *treatment, contingency* and *measurement phase* was also highly significant. Otherwise, as Table 2.12 shows, there was a suggestion of interactive effects in some of the higher-order terms featuring *probe value,* however, taken as a whole (i.e. considering the multiplicity of effects examined), these were thought not strong enough to merit inclusion in the model.

Therefore, the *treatment*measurement phase* and *treatment*contingency* measurement phase* terms remained (together with any lower-order terms necessary to satisfy marginality), and the predicted probability of pressing the lever associated with 2 pellet of food was estimated from this model, with the *subject*-level variance removed (i.e. from the model summarised in Table 2.14 below, and specified in Equation 0.9, in the Appendix; Table 0.2 and Figure 0.14, also in the Appendix, present the results of a Hosmer-Lemeshow goodness-of-fit test for this model, which was satisfactory). Figure 2.23 and Figure 2.24 plot these predictions, for each of the *measurement phase / treatment / contingency* groups.

Parameter		Coefficient estimate (with SE)	Wald (x2)	Df	*P*
*Treatment*probe value*	Fixed	0.187 (0.403)	0.216	1	0.642
Treatment(probe value)^2*	Fixed	0.258 (0.517)	0.250	1	0.617
Treatment(probe value)^3*	Fixed	-0.222 (0.704)	0.099	1	0.753
Treatment Contingency*	Fixed	0.195 (0.264)	0.545	1	0.460
*Treatment*contingency* probe value*	Fixed	-1.824 (0.788)	5.351	1	0.021 *
*Treatment*contingency* (probe value)^2*	Fixed	-1.471 (1.039)	2.003	1	0.157
*Treatment*contingency* (probe value)^3*	Fixed	2.305 (1.547)	2.219	1	0.136
*Treatment*measurement phase*probe value*	Fixed	0.123 (0.388)	0.101	1	0.751
*Treatment*measurement phase*(probe value)^2*	Fixed	-0.965 (0.391)	6.095	1	0.014 *
*Treatment*measurement phase*(probe value)^3*	Fixed	0.095 (0.895)	0.011	1	0.916
*Treatment*Contingency *Measurement phase*	Fixed	-0.648 (0.211)	9.422	1	0.002 **

Table 2.12 As Table 2.6, but derived from a modifed model, as specified in Equation 0.8 in the Appendix, but with the addition of all terms necessary to respect marginality.

Parameter		Coefficient estimate (with SE)	Wald (x2)	Df	*P*
Treatment	Fixed	-0.228 (0.144)	2.506	1	0.113
*Treatment*Measurement phase (Phase 1 =ref. category)*	Fixed	0.326 (0.105)	9.595	1	0.002 **
Treatment	Fixed	0.098 (0.143)	0.465	1	0.495
*Treatment*Measurement phase (Phase 2 = ref. category)*	Fixed	-0.326 (0.105)	9.595	1	0.002 **

Table 2.13 As Table 2.6, but derived from a modifed model, as specified in Equation 0.7 (with the addition of the above terms).

The significant *contingency*probe value* interactions we found in the model are reflected in the different shape of the response curves across *contingency* group, with a steeper transition between the areas of highest and lowest peak responding for the *2kHz=2pell contingency* group, and a greater difference between the areas of highest and lowest peak responding in this *contingency* group too.

It is also very apparent from these plots that the main change across *measurement phase* occurs for the *control / 2kHz=2pell* group, whose probability of pressing the lever associated with 2 pellets of food is markedly lower in *Phase 2*. The responding of the other groups remains very similar across *measurement phase*, and so it is very likely that the responding of the *control / 2kHz=2pell* group underlies the significant *treatment*measurement phase*contingency* interaction, as well as the lower order interactions between *treatment* and *measurement phase*, and also *measurement phase* and *contingency*. The predicted response pattern of the *control / 2kHz=2pell* group in *Phase 2* is similar to that of the other *treatment* group in that *contingency* (i.e. *UHT / 2kHz=2pell*) which varies very little from one *measurement phase* to the other; therefore the responding of the *control / 2kHz=2pell* group in *Phase 1* is comparatively unusual.

Parameter		Coefficient estimate (with SE)	Wald (x2)	Df	*P*
Intercept	Random at *subject* level	0.081 (0.036)	4.995	1	0.025 *
Probe value	Random at *subject* level	0.308 (0.134)	5.278	1	0.022 *
	Fixed	5.999 (0.304)	390.616	1	<0.001 **
(Probe value)^2	Random at *subject* level	0.737 (0.375)	3.884	1	0.049 *
	Fixed	0.913 (0.393)	5.413	1	0.020 *
(Probe value)^3	Fixed	-8.073 (0.570)	200.894	1	<0.001 **
Contingency	Fixed	0.306 (0.222)	1.912	1	0.167
*Contingency*Probe value*	Fixed	-2.027 (0.414)	24.023	1	<0.001 **
Contingency(Probe value)*^{2}	Fixed	-2.684 (0.551)	23.759	1	<0.001 **
Contingency(Probe value)*^{3}	Fixed	3.313 (0.771)	18.449	1	<0.001 **
Measurement phase	Fixed	-0.673 (0.107)	39.366	1	<0.001 **
*Measurement phase *Contingency*	Fixed	0.704 (0.148)	22.542	1	<0.001 **
Treatment	Fixed	-0.525 (0.211)	6.167	1	0.013 *
*Treatment*Contingency*	Fixed	0.536 (0.286)	3.519	1	0.061
*Treatment*Measurement phase*	Fixed	0.673 (0.155)	18.925	1	<0.001 **
*Treatment*Contingency *Measurement phase*	Fixed	-0.648 (0.211)	9.422	1	0.002 **

Table 2.14 The coefficient estimates, with standard error, together with Wald tests and their significance, from the final model with the key terms of interest: *probe value*, *contingency*, *measurement phase* and *treatment* (as specified in Equation 0.9, in the Appendix).

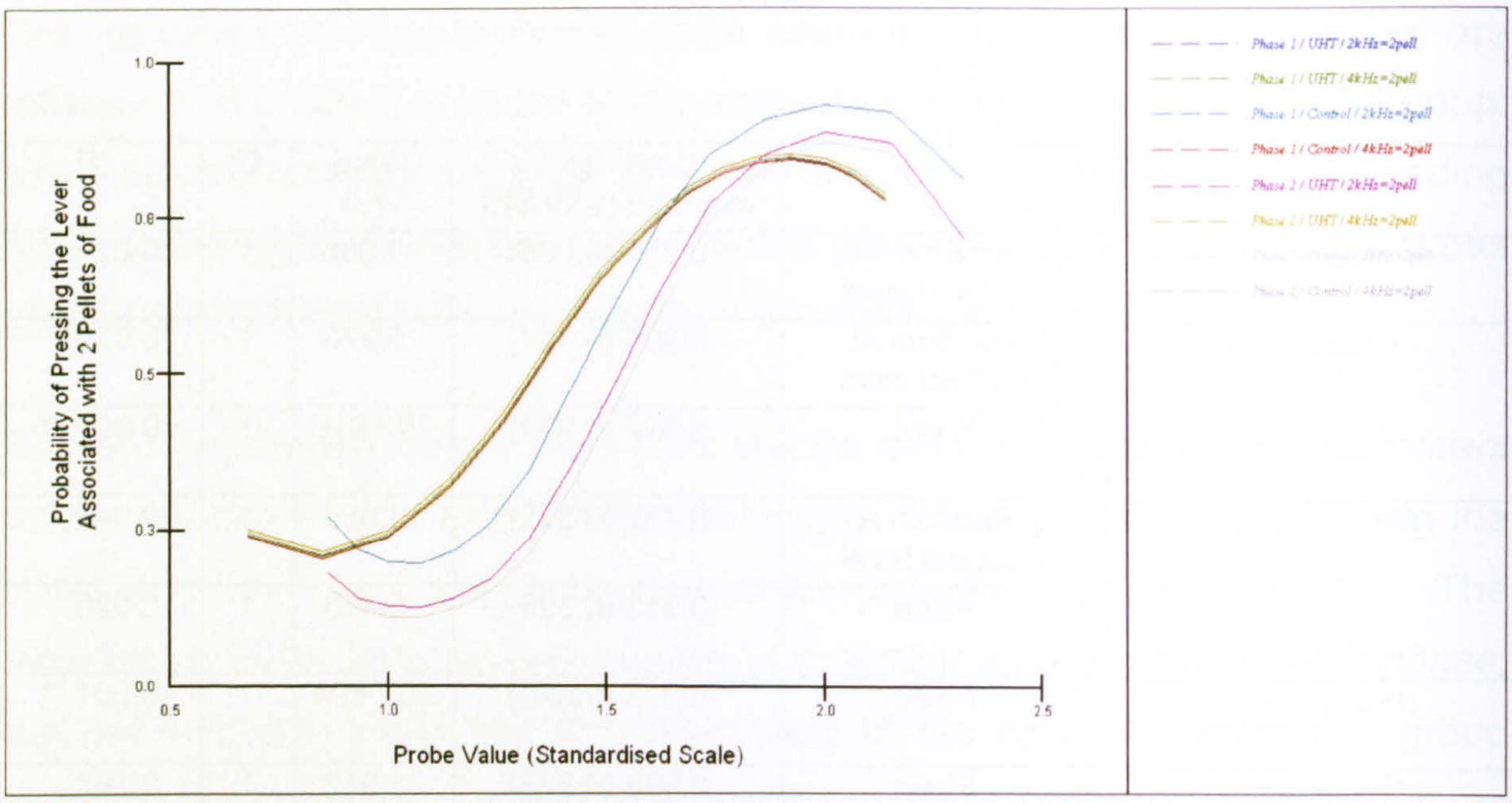

Figure 2.23 The probability of pressing the lever associated with 2 pellets of food (as opposed to the lever associated with 1 pellet of food) across different values of the probe stimuli (a standardised scale, with *1.0* corresponding to the reference tone associated with 1 pellet of food, and *2.0* corresponding to the reference tone associated with 2 pellets of food), by *measurement phase / treatment / contingency* group. These predicted lines were generated from the model and estimates in Equation 0.9. Although it is not clear from this graph, the line for the *Phase 1 / UHT / 2kHz=2pell* group is behind that of the same group in the 2nd *measurement phase* (i.e. behind *Phase 2 / UHT / 2kHz=2pell)*.

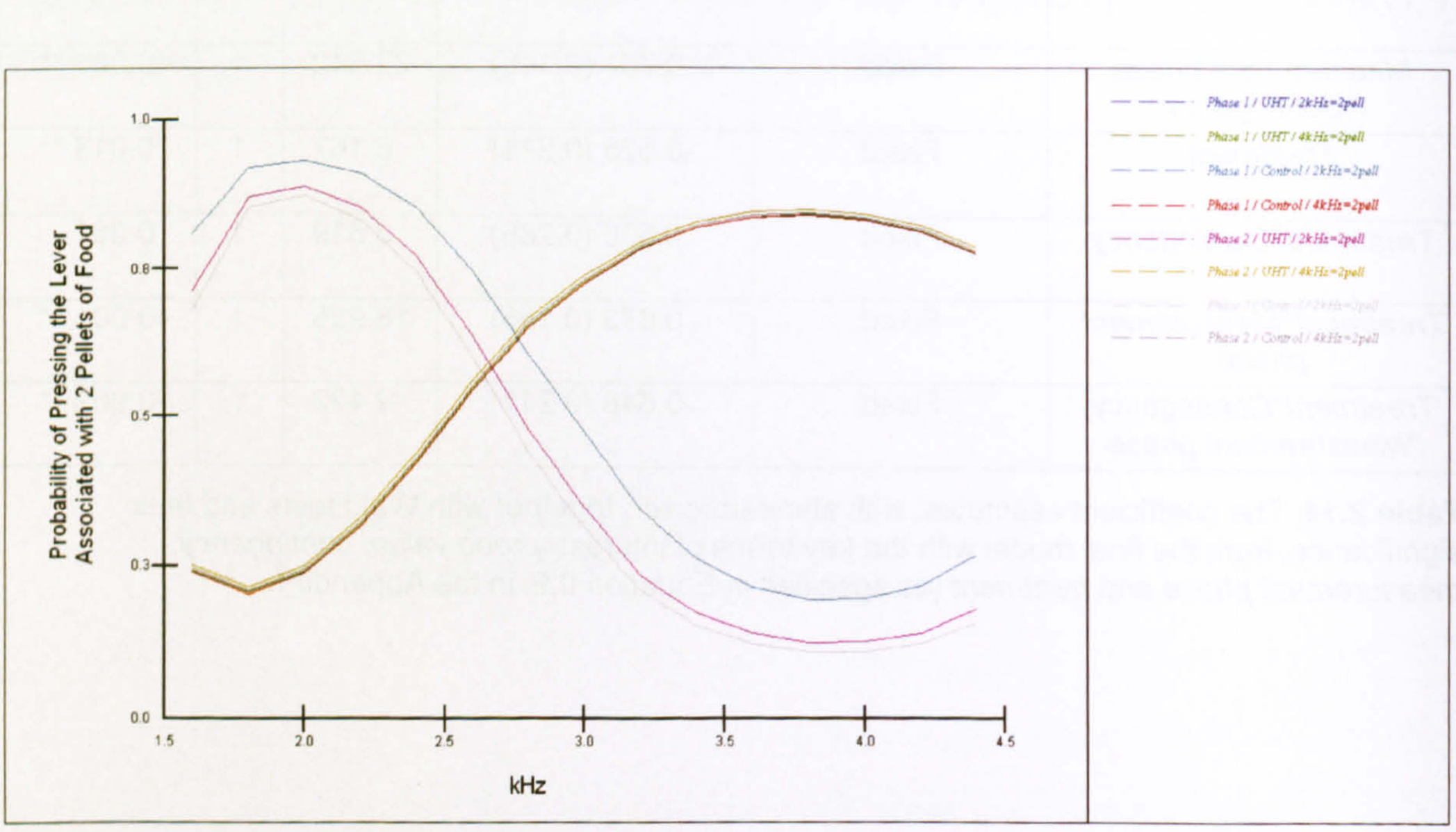

Figure 2.24 As Figure 2.23, but with kHz on the x-axis.

If we go back to the predicted response curves we plotted from an earlier model, in Figure 2.21 and Figure 2.22, it is interesting to note that the pattern of responding by *Subject 6* and *Subject 8*, two of the four rats in the *control / 2kHz=2pell* group, is

somewhat different to that of the other *subjects*. If we generate these predictions again, but from equations which only model the data from each of the two *measurement phases*, we see that these two *subjects* have a particularly high probability of pressing the '2-pellet' lever in *phase 1* (Figure 2.25), but that this difference is considerably diminished in *phase 2* (Figure 2.26; here, we only plot the data with *kHz* on the x-axis, rather than *probe value* as well, as it is a little easier to distinguish the predicted response lines).

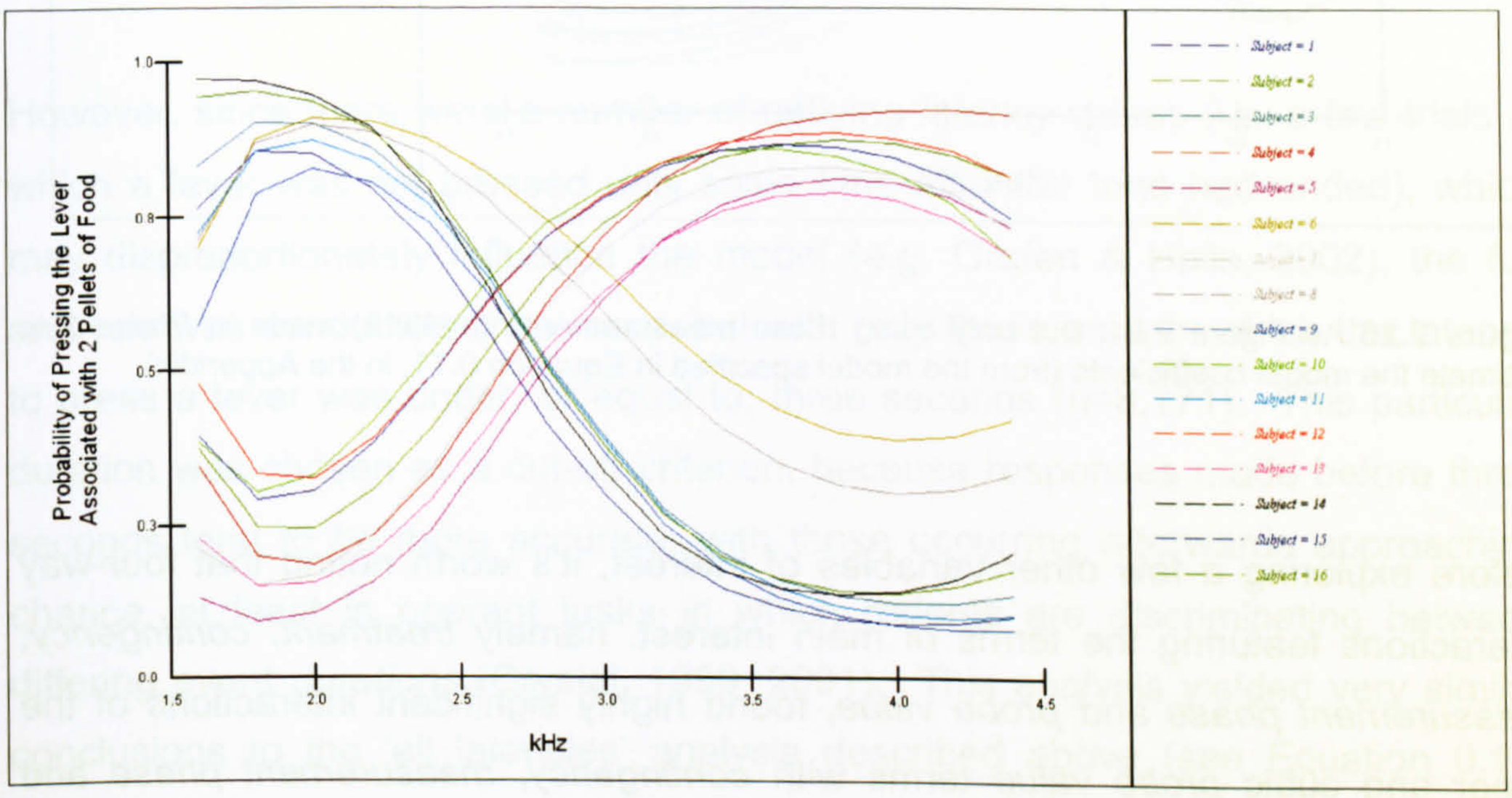

Figure 2.25 As Figure 2.22, but with only those measurements (n=4307) made in *Phase 1* (i.e. at baseline) used to estimate the model coefficients (from the model specified in Equation 0.10, in the Appendix).

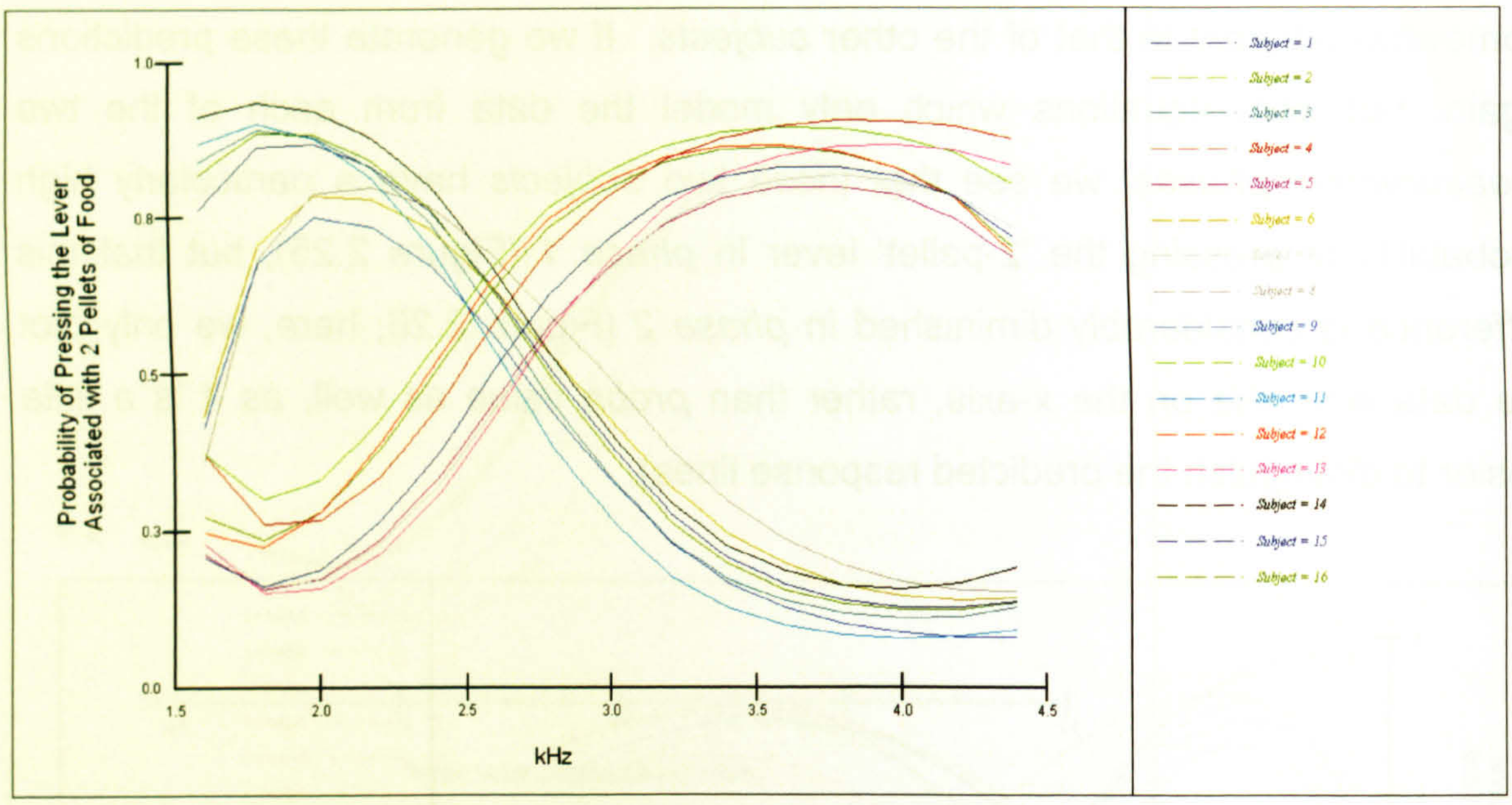

Figure 2.26 As Figure 2.21, but only using those measurements (n=4319) made in *Phase 2* to estimate the model coefficients (from the model specified in Equation 0.11, in the Appendix).

Before exploring a few other variables of interest, it's worth noting that four-way interactions featuring the terms of main interest, namely *treatment, contingency, measurement phase* and *probe value,* found highly significant interactions of the linear and cubic *probe value* terms with *contingency, measurement phase* and *treatment* (for reference, see Table 0.3, Equation 0.12, and Figure 0.15 to Figure 0.16, in the Appendix). Clearly, the nature of this interaction would be very complex to distil, and our inspection of the predicted responses plotted from this model found no indication of any differences which would support our experimental hypotheses. More generally, as discussed earlier, such an elaborate equation (necessarily so, to respect marginality, e.g. Grafen & Hails, 2002) risks over-fitting the model: i.e. re-describing the observed data, with little gain in predictive or explanatory power (e.g. Quinn & Keough, 2002).

These higher-order terms were deleted, returning the model to that presented in Table 2.14 (i.e. Equation 0.9, in the Appendix). *Latency* to press the lever was then added as a main effect, and it approached significance at the 0.05 level, with the coefficient value indicating that rats took longer to press the lever associated with 2 pellets of food (see Table 2.15, and Equation 0.13 in the Appendix); two-way

interactions of *latency* with the *probe value* terms were also added, but the model would not converge, and so these were removed.

Parameter		Coefficient estimate (with SE)	Wald (χ2)	Df	*P*
Latency	Fixed	0.009 (0.005)	3.124	1	0.078

Table 2.15 As Table 2.6, but derived from a modifed model, as specified in Equation 0.13, in the Appendix.

However, since there were a number of outlying *latency* values (i.e. a few trials in which a lever was not pressed until some time after the tone had ended), which may disproportionately influence the model (e.g. Grafen & Hails, 2002), the full analysis was conducted once more, but with only those trials in which the latency to press a lever was under, or equal to, three seconds (n=8,171). This particular duration was chosen as a cut-off criterion, because responses made before three seconds tend to be more accurate, with those occurring afterwards approaching chance, at least in operant tasks in which rodents are discriminating between differing event *durations* (Crystal, 1999, 2001). This analysis yielded very similar conclusions to the 'all latencies' analysis described above (see Equation 0.14, Figure 0.17 and Figure 0.18, and also Table 0.4; all in the Appendix), but when *latency* was added to this 'under 3 seconds' model, it was now highly significant, indicating a significantly higher probability of pressing the lever associated with 2 pellets of food as *latency* increased (see Table 2.16, and Equation 0.15 in the Appendix).

Parameter		Coefficient estimate (with SE)	Wald (χ2)	Df	*P*
Latency	Fixed	0.197 (0.053)	13.950	1	<0.001 **

Table 2.16 The coefficient estimates, with standard error, of certain fixed parts of the model (limited to responses of a latency no greater than 3 seconds), together with Wald test statistics and their significance (from the model specified in Equation 0.15, in the Appendix).

Two-way interactions of *latency* with *probe value* were also explored, of which the interactions with the linear and cubic terms were highly significant (see Table 2.17,

and Equation 0.16 in the Appendix). *Latency* is clearly an important variable in the analysis, and investigations of its effect are pursued further when it is modelled as the response (*y*) variable (see below).

Parameter		Coefficient estimate (with SE)	Wald (χ2)	Df	*P*
Latency	Fixed	0.062 (0.068)	0.842	1	0.359
*Latency*Probe Value*	Fixed	-2.915 (0.234)	155.243	1	<0.001 **
Latency(Probe Value)^2*	Fixed	0.291 (0.241)	1.457	1	0.227
Latency(Probe Value)^3*	Fixed	3.245 (0.514)	39.935	1	<0.001 **

Table 2.17 The coefficient estimates, with standard error, of certain fixed parts of the model (limited to responses of a latency no greater than 3 seconds), together with Wald test statistics and their significance (from the model specified in Equation 0.16, in the Appendix).

Finally, the possibility of adding other terms of interest to these models was investigated. Correlation matrices indicated that *latency* was correlated with a number of other variables, such as *session number (i.e. 1-6)*, performance in recent training sessions (e.g. *percentage of 2-pellet trials correct; percentage of 1-pellet trials correct*), *trial number (1-156)*, *within-session (pseudorandomised) block (1-3)*, *time since session started (in seconds)*, etc. Therefore, whilst these terms are of some general interest, introducing them alongside *latency* would lead to problems of multiple collinearity (e.g. Grafen & Hails, 2002; Quinn & Keough, 2002); i.e. a number of terms share information with *latency*, and the most parsimonious strategy is to keep the latter term in the model alone. The *position of the 2-pellet lever* (i.e. whether the lever, on which correct responses were associated with reinforcement with two pellets of food, was on the left or right of the food hopper) was part of the orthogonal design of the experiment, however, and so the contribution of this variable was explored. *Left position* was assigned the reference category, with a value of *'0'*, whilst *right position* had a value of *'1'*. Interestingly, this predictor variable was highly significant. In both models (i.e. with 'all latencies', or just with those in which a lever press was made 'within 3 seconds': Equation 0.17 and Equation 0.18 in the Appendix, respectively), the probability of pressing the lever associated with 2 pellets of food was significantly

greater when it was on the right of the food hopper (and the 1-pellet lever was on the left, as opposed to *vice versa*). In addition, this variable explained a great deal of the *subject*-level variance in the overall probability of lever choice, although the unexplained *subject*-level variance in lever choice across *probe value* remained significant. Table 2.18 shows these coefficient estimates, which (for brevity and consistency's sake) are taken from the 'all latencies' model. The high significance of the *'Position of the 2-pellet lever'* term persisted after *Room* (i.e. the room in which the rat's were trained and tested: either Room *A* or *B*), and a two-way interaction of *Room* with *Position of the 2-pellet lever,* were added, neither of which were significant (see Equation 0.19 and Equation 0.20, in the Appendix, for the respective models). Finally, it's interesting to note that the two *subjects* who had a considerable bias towards pressing the '2-pellet' lever in *measurement phase 1* (*subjects 6* and *8;* see Figure 2.25), both had the '2-pellet' lever assigned to their right-hand side.

Parameter		Coefficient estimate (with SE)	Wald (χ2)	Df	*P*
Intercept	Random at *subject* level	0.012 (0.012)	1.067	1	0.302
Probe value	Random at *subject* level	0.303 (0.131)	5.379	1	0.020 *
	Fixed	6.009 (0.302)	396.513	1	<0.001 **
(Probe value)^2	Random at *subject* level	0.729 (0.370)	3.874	1	0.049 *
	Fixed	0.921 (0.391)	5.547	1	0.019 *
Position of the 2-pellet lever	Fixed	0.534 (0.063)	71.852	1	<0.001 **

Table 2.18 As Table 2.6, but derived from a modifed model, as specified in Equation 0.17, in the Appendix.

Summary of single-frequency probe lever choice analyses

The multilevel analysis in MLwiN found a range of significant interactions, including *measurement phase* with *treatment* (indicating an 'optimistic' shift in the responding of the *UHT* rats, compared to the *Control* group), *measurement phase* with *contingency,* and a significant three-way interaction between all three terms. Looking more closely at the performance of individual rats in the two *measurement phases,* two *subjects*, in the *Control / 2kHz=2pell* group, with an unusually high

probability of pressing the '2-pellet' lever in *phase 1* were likely to have made a particularly large contribution to these findings, and also to the highly-significant contribution of a term pertaining to the position (left or right) of the '2-pellet' lever.

Otherwise, the pattern of responding of the *4kHz=2pell contingency* group described a shallower curve across *probe value*, with a smaller difference between the points of maximum and minimum responding on the '2-pellet' lever, than the *2kHz=2pell contingency* group. More generally, the analysis indicated that the *2kHz=2pell contingency* group were overall more 'pessimistic': i.e. overall less likely to press the '2-pellet' lever. Presses on the '2-pellet' lever were overall slower than presses on the '1-pellet' lever, although further investigations of *latency* were left to subsequent analyses, described in the relevant section below.

In general, these findings were in good agreement with the results from a preceding analysis (see Appendix C), in which repeated-measures ANOVAS were conducted on various aspects of fitted probit functions. For example, the latter indicated relatively little change in the point of bisection (i.e. the probe value at which there was an estimated even chance of pressing either lever) across *measurement phase,* with the notable exception of subjects in the *Control / 2kHz=2pell* group, who were considerably more 'pessimistic' in the second *measurement phase* compared to the first. In addition, the analyses of the fitted probit functions concurred closely with the fitted curves derived in MLwiN (e.g. see Figure 2.23), with the *2kHz=2pell* group more 'pessimistic' at the point of bisection, and also with regard to responding to the intermediate *probe value* closest to the '1-pellet' reference tone, whilst there was no significant main effect of *contingency* with respect to responding to the intermediate *probe value* closest to the '2-pellet' reference tone.

Single-frequency test sessions: latency

Figure 2.27 and Figure 2.28 plot the mean latency to press the lever associated with 2 pellets of food, across kHz, for the *2kHz=2pell* and *4kHz=2pell contingency* groups, respectively, whilst Figure 2.29 and Figure 2.30 plot the mean latency to press the lever associated with 1 pellet of food, across kHz, for the *2kHz=2pell* and

4kHz=2pell contingency groups, respectively. As discussed on page 70, the *latency* data was analysed in a multi-level model using the statistical software package MLwiN, following an earlier analysis using repeated-measures ANOVAs in SPSS (see Appendix C); the former analysis is described below, and concludes with a summary section comparing the results with the latter.

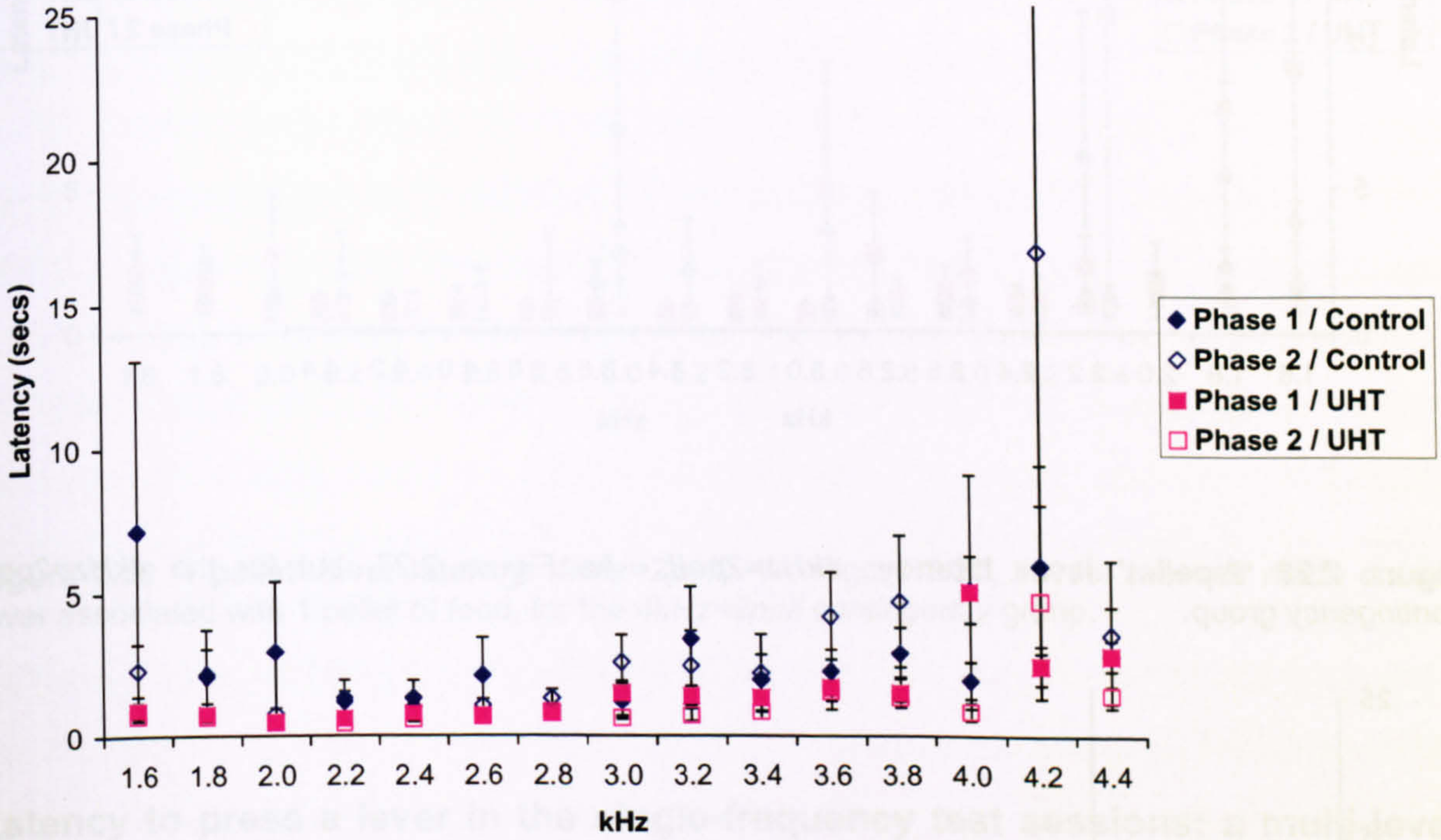

Figure 2.27 '2-pellet' lever latency: *2kHz=2pell*. The mean latency, in seconds, for the *2kHz=2pell contingency* group to press the lever associated with 2 pellets of food, by *measurement phase / treatment* group, across kHz (+/- 1 SEM. N.B. the means and SEM are derived from data summarised at the *subject*-level (note: if, in one of the *measurement phases,* a *subject* never pressed this particular lever following one of the kHz tones, that *subject* did not contribute to the corresponding summary datapoint[69]); in addition, the data pertaining to the 'reference tones' (i.e. 2kHz and 4kHz) are taken from the non-reinforced trials only).

[69] An alternative would have been to enter a maximum 'timed-out' latency, but since each trial terminated following a lever press (rather than being timed-out), it is not clear what value any such maximum latency would take.

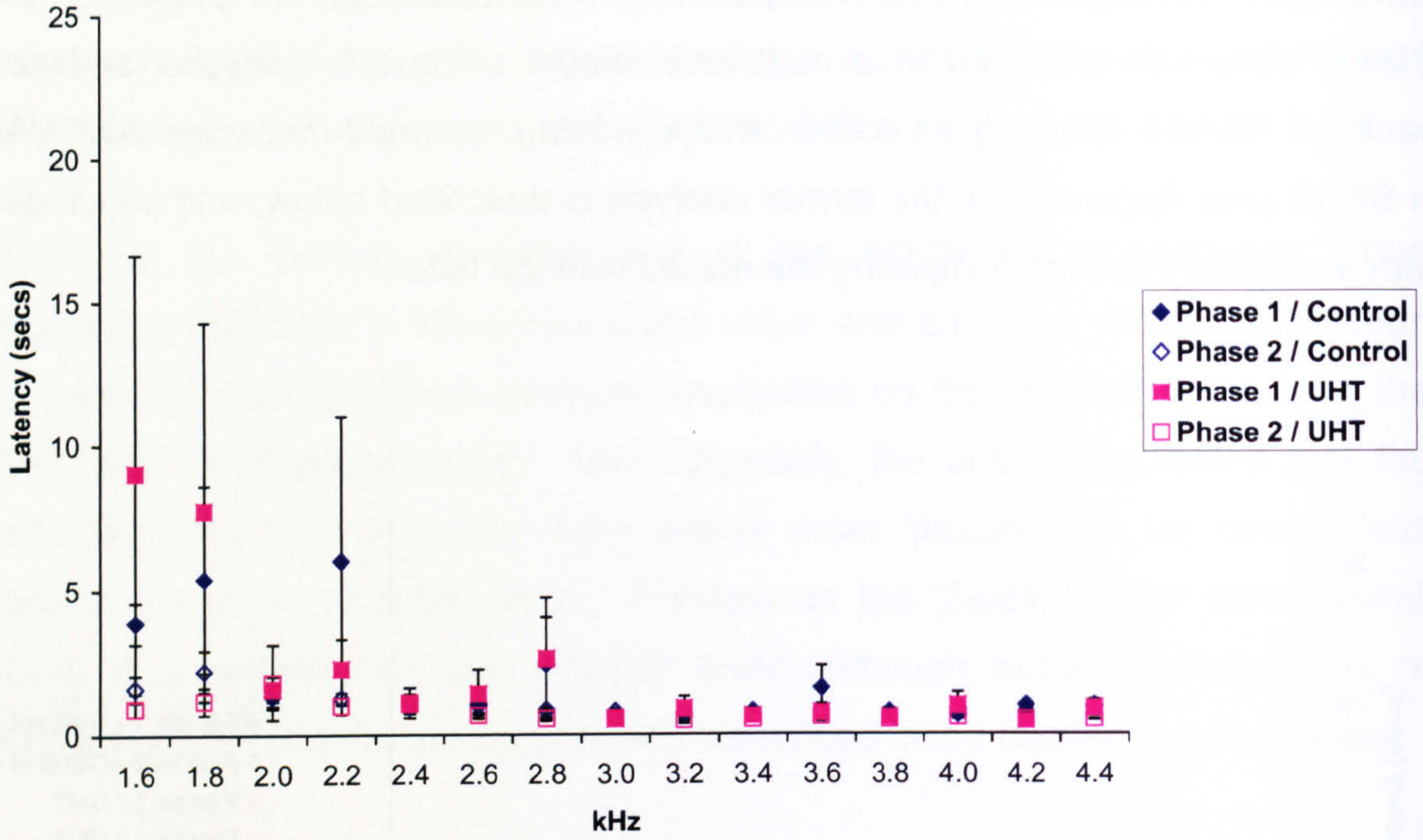

Figure 2.28 '2-pellet' lever latency: *4kHz=2pell*. As Figure 2.27, but for the *4kHz=2pell* *contingency* group.

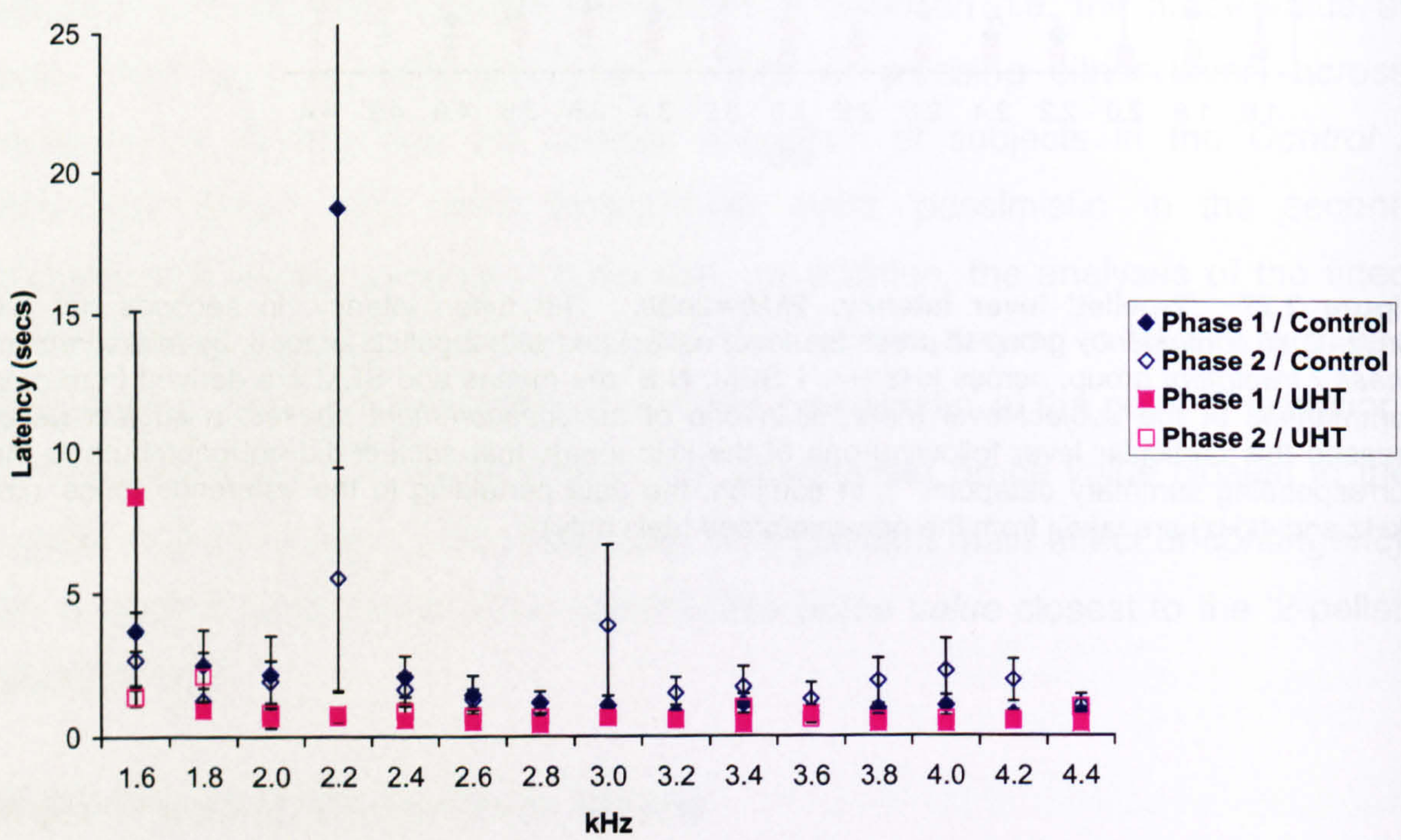

Figure 2.29 '1-pellet' lever latency: *2kHz=2pell*. As Figure 2.27, but for presses made on the lever associated with 1 pellet of food.

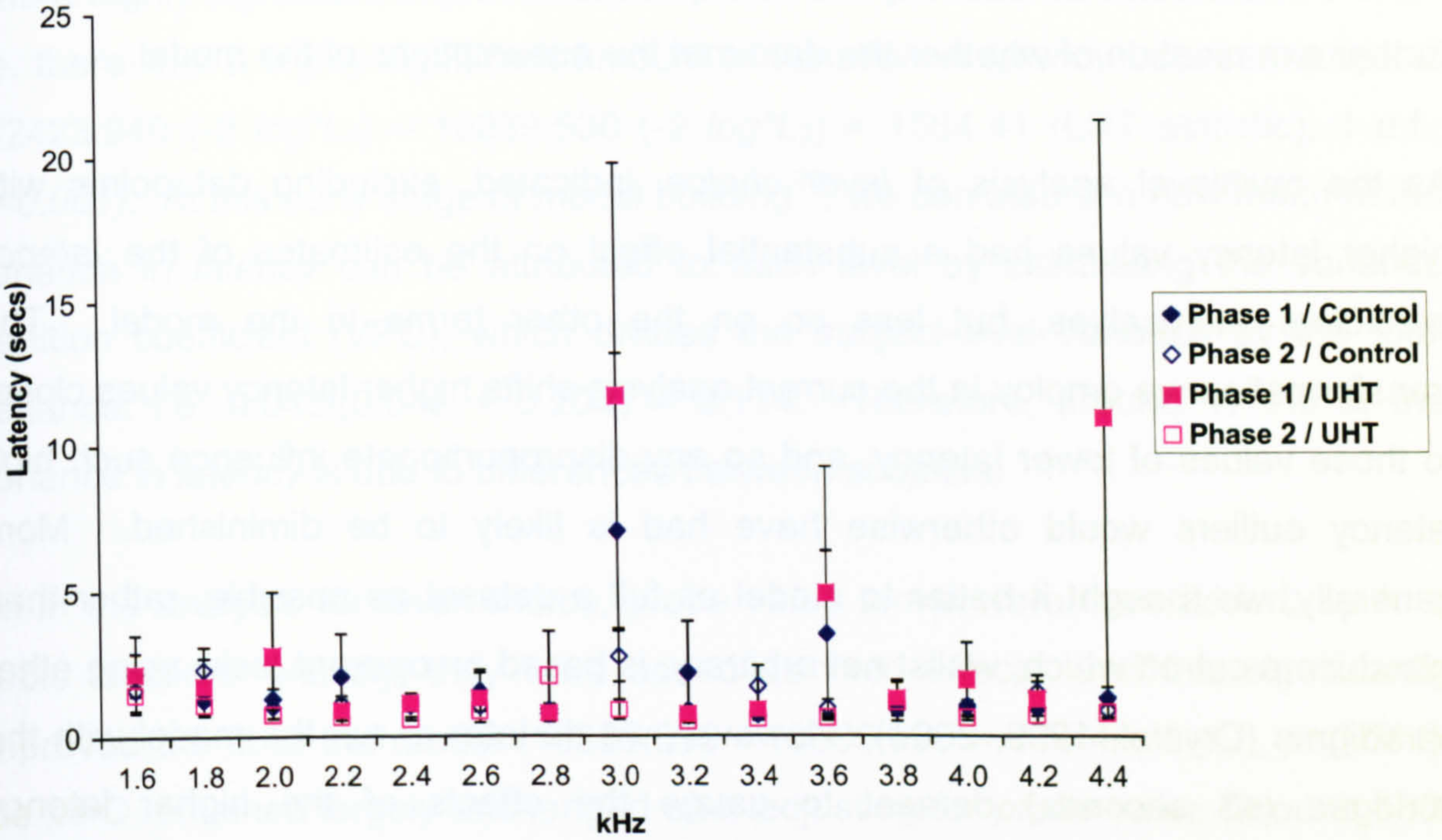

Figure 2.30 '1-pellet' lever latency: *4kHz=2pell*. As Figure 2.27, but for presses made on the lever associated with 1 pellet of food, for the *4kHz=2pell contingency* group.

Latency to press a lever in the single-frequency test sessions: a multi-level general linear model in *MLwiN*

As discussed on p.74, the *latency* to press a lever in the single-frequency probe sessions was analysed in a multilevel model using MLwiN (of course, it was introduced as a predictor (x) variable in the multilevel analysis of *lever choice*, above, but here it is modelled as the response (y) variable).

Since our inspections of the raw *latency* data revealed a marked positive skew, some form of transformation was likely to be necessary to meet the assumptions of a general linear model. After exploring various models and accompanying residuals, a negative reciprocal root transformation was chosen[70] with seven

[70] i.e. $-1/\sqrt{y}$ (the minus sign maintains the same order in the data). Our initial testing of the assumptions found a log-transformation did not sufficiently normalise the data (i.e. was too 'weak'), and a negative inverse transformation increased the variance of the smaller predicted (fitted) values too greatly (i.e. was too 'strong'); the negative reciprocal root, intermediate to these two in 'strength' of transformation (e.g. Emerson, 1991), was therefore thought a good choice.

outliers of the shortest latency omitted (all ≤ 0.12 seconds)[71]. See later for a further examination of whether the data met the assumptions of the model.

As the multilevel analysis of *lever choice*, indicated, excluding datapoints with higher latency values had a substantial effect on the estimates of the *latency* predictors themselves, but less so on the other terms in the model. The transformation we employ in the current analysis shifts higher latency values closer to those values of lower latency, and so any disproportionate influence such high latency outliers would otherwise have had is likely to be diminished. More generally, we thought it better to model as full a dataset as possible, rather than introduce a cut-off which, whilst not arbitrary, is based on operant tasks using other paradigms (Crystal, 1999, 2001). However, we do later re-run the model with the abridged (≤3 seconds) dataset to gauge the effects of the higher latency datapoints.

The hierarchy of the dataset was defined in the same way as in the multilevel analysis of *lever choice:* i.e. with *trial* (n=8619) at Level 1, nested within *subject* (n=16) at Level 2.

The value of *-2 log*likelihood* (see p. 76) was taken from a single-level model with only an intercept term (Equation 0.21 in the Appendix shows the output from MLwiN: the relevant statistic is listed at the bottom)[72]. The intercept was then allowed to randomly vary at the *subject*-level (i.e. a simple 'random intercept' model was fitted; Equation 0.22 in the Appendix), and the value of *-2 log*likelihood* was taken from this model. An LRT statistic was then obtained by subtracting one of these values from the other, and the significance of this statistic was obtained by reference to a χ^2 distribution with one degree of freedom. This indicated that there

[71] The transformation used in this analysis shifts very small values (e.g. ≤0.12) far away from datapoints of larger value (as would a negative inverse transformation). When examining whether the assumptions of the model were satisfied, these appeared as far outliers on the plots of the standardised residuals we inspected. Since there were so few points with a latency under or equal to 0.12 seconds (7 out of a dataset of 8626), which did not form a separate 'population' in the series (i.e. latencies of 0.14, 0.16, 0.18, etc. seconds, remained in the dataset), it was thought very unlikely that excluding these points would change the dataset in a manner which would misrepresent the biological significance of the data, and so they were simply omitted so that the assumptions of the model could be better satisfied.

[72] As in the multilevel analysis of *lever choice*, it's not necessary to refer to all the equations provided in the Appendix – they are simply for reference, should the reader wish to study other aspects of the model not reported in the main text.

was a highly-significant improvement in the fit of the model to the observed data: i.e. there was a highly-significant amount of variation in *latency* between *subjects* (12423.940 (*-2 log*L_1*) – 10839.530 (*-2 log*L_2*) = 1584.41 (LRT statistic), 1 d.f., *p<0.001).* At this early stage of model building[73], we can also see how much of the variation in *latency* can be attributed to each level by calculating the variance partition coefficient (VPC), which divides the *subject*-level variance by the total variance: i.e. 0.043/(0.043 + 0.204) = 0.174. Therefore, around 17.4% of the variance in *latency* is due to differences between *subjects*.

As in the analysis for *lever choice*, *probe value* terms were then added, up to a cubic function. Initially, they were added as fixed effects, and this significantly improved the fit of the model (10839.530 - 10665.090 = 174.44, 3 d.f., *p<0.001*); the VPC remained largely unchanged after adjusting for *probe value* (around 18% of variance is attributable to differences between *subjects*; see Equation 0.23). The coefficients of the linear and quadratic terms were then allowed to vary across *subject* (Equation 0.24; the cubic term remained a fixed effect, as in the analysis for *lever choice*), which further improved the fit of the model (10665.090 – 10544.000 = 121.09, 5 d.f., p<0.001) [74]: i.e. there was a significant amount of *subject*-level variance in *latency* across *probe value*.

Lever choice (i.e. whether the lever associated with 2 pellets of food was pressed or not) was then added as a fixed effect, and the significance of the *likelihood ratio test* indicated this modification improved the model's fit (10544.000 – 10500.380 = 43.62, 1 d.f., p<0.001; Equation 0.25, in the Appendix), indicating a longer *latency* when a response on the lever associated with 2 pellets of food was chosen (as in the multilevel analysis with *lever choice* as the response (y) variable, described above).

---

[73] The variance partition coefficient is less appropriate once a 'random slope' is introduced (e.g. see Centre for Multilevel Modelling (CMM)'s online 'frequently asked questions': http://www.cmm.bristol.ac.uk/MLwiN/tech-support/support-faqs/pval.shtml).

74 This difference in the degrees of freedom is due to the extra variances and covariances produced when these additional coefficents are allowed to vary at the *subject* level (see the covariance matrix at the bottom of Equation 0.24) (Rasbash et al., 2005).

The shape of the regression curve was then allowed to differ across *lever choice* (i.e. the *latency* pattern was allowed to differ across *probe value*, depending on whether the '1-pellet' or '2-pellet' lever was pressed), with the addition of two-way interactions featuring *lever choice* and the *probe value* terms. Again, this resulted in a considerably better fit of the model to the observed data (10500.380 – 9960.326 = 540.054, 3 d.f., p<0.001; Equation 0.26, in the Appendix).

Contingency, which was then added as a main effect, was not a significant term (apparent, for example, by comparing the LRT statistics (9960.326 – 9960.324 = 0.002, 1 d.f., p=0.964), or conducting a Normal distribution test on the ratio between the estimated coefficient of *contingency*, and its standard error (0.004/0.100 = 0.04, p=0.484); Equation 0.27, in the Appendix), indicating no overall difference in *latency* between the two *contingency* groups.

Interactions of *contingency* with the *probe value* terms and with *lever choice* were then explored up to a three-way interaction. The full model (i.e. with the three-way interaction and all lower-order terms necessary to respect marginality) resulted in a highly-significant improvement in the fit of the model (9960.324 – 9814.120 = 146.204, 7 d.f., p<0.001; Equation 0.28, in the Appendix). As discussed in the multilevel analysis of *lever choice* (above), fitting high-order interactions to a model can risk over-fitting: compromising explanatory, or predictive, power with an over-elaborate equation. However, in this instance, there are theoretical reasons to include these interactions (e.g. Aiken & West, 1991): for example, it's reasonable to expect the two *contingency* groups to differ in the speed with which they press one or other of the levers (e.g. perhaps on account of one of the reference tones (i.e. 2kHz or 4kHz) being more salient than the other), to therefore differ across *probe value*, and for there to be an interaction between all three. Therefore, it was thought best to preserve this three-way interaction.

Similarly, *measurement phase*, when added solely as a main effect, made little contribution to the model (9814.120 – 9811.991 = 2.129, 1 d.f., p=0.145; Equation 0.29, in the Appendix), indicating there was no overall effect of *measurement phase* on the speed with which a lever response was made. However, the fit of the model was improved when *measurement phase* was interacted with *lever choice*

(Equation 0.30; 9811.991 – 9802.132 = 9.859, 1 d.f., p=0.002), and an inspection of the coefficients revealed that the *subjects* were overall quicker to press the '2-pellet' lever, and slower to press the '1-pellet' lever, in the second *measurement phase,* compared to the first *measurement phase.* There was also a significant interaction (in a variety of models we investigated) between *measurement phase* and *contingency*, and an inspection of the coefficients revealed that the *2kHz=2pell* group were slower, and the *4kHz=2pell* group were faster, to record a lever press response in the second *measurement phase,* compared to the first (e.g. 9802.132 – 9788.931 = 13.201, 1 d.f., p<0.001; Equation 0.31). The interaction between *measurement phase* and the *probe value* terms did not improve the fit of the model (again, this was explored in a variety of equations), indicating there was no change in the shape of the regression curve between the two *measurement phases* (e.g. 9788.931 – 9783.341 = 5.590, 3 d.f., p=0.133; Equation 0.32, in the Appendix). Rather than fit higher-order interactions featuring *measurement phase,* in order to preserve some parsimony in the model, the higher-order terms were limited to the two-way interactions explored above, of which *measurement phase*lever choice* and *measurement phase*contingency* remained.

The addition of the final term of main interest, *treatment*, significantly improved the fit of the model, indicating that rats in the *UHT* group had a significantly shorter *latency*: i.e. made significantly quicker lever press responses (9788.931 – 9781.123 = 7.808, 1 d.f., p=0.005; Equation 0.33). *Treatment*Measurement phase* further improved the fit of the model to the observed data (9781.123 – 9772.200 = 8.923, 1 d.f., p=0.003; Equation 0.34), and inspection of the coefficients indicated that the *UHT* group were quicker to record a lever press response, compared to the *control* group, in both *measurement phases,* but this difference was greater in *measurement phase 2:* i.e. the *latency* of the *Control* group to record a lever press became greater across *measurement phase* (i.e. they got slower), whereas the *latency* of the *UHT* group to record a lever press became slightly smaller across *measurement phase* (i.e. they got slightly faster).

A variety of other interactions, featuring *treatment,* were explored, in a variety of models: a number of the lower-order interactions made a non-significant contribution to the model (e.g. *Treatment*Contingency; Treatment*lever choice;*

*Treatment*probe value; Treatment*contingency*measurement phase; Treatment*measurement phase*lever choice*), whilst models fitted up to higher-order interactions (e.g. *Treatment*measurement phase*lever choice*probe value*, etc.*)* did significantly improve the fit. However, whilst some of these higher-order interactions are of theoretical interest, it's rather hard to distil meaning from such elaborate models; therefore we leave some of this closer analysis to two subsidiary investigations, described later in this section, which only model the data from one lever.

These interactions were removed, leaving only *treatment* and *treatment*measurement phase* (Equation 0.34, in the Appendix), and predictions were generated from this equation. Figure 2.31 to Figure 2.34 chart the results[75]. The different patterns of latency across *probe value* for each of the *contingency* groups is apparent in the plots, as is the main effect of *treatment,* with the rats in the *UHT* group recording a lever press consistently faster than the rats in the *Control* group. The plots of the residuals, in the Appendix (see Figure 0.19 to Figure 0.22), suggest the model's assumptions are reasonably well upheld. In addition, the model was re-run, but only modelling those datapoints of an (untransformed) latency under or equal to 3 seconds (n=8,164): as Equation 0.35, in the Appendix, shows, whilst there are some changes to the coefficient estimates (as would be expected when modelling an abridged dataset), the interpretations from the model remain very similar.

[75] Backtransformation = $1/(y^2)$

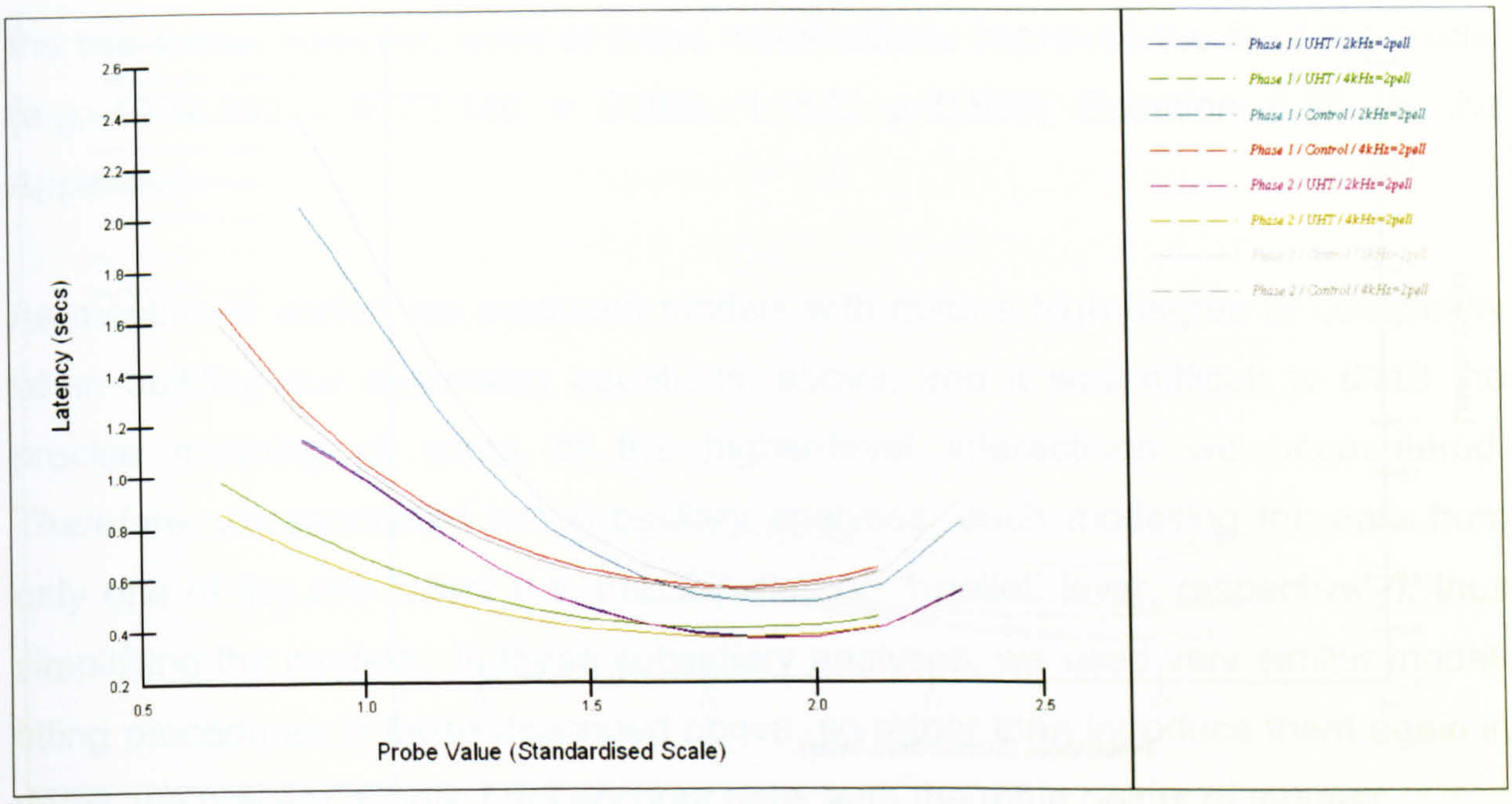

Figure 2.31 The latency (in seconds) to press the lever associated with 2 pellets of food (as opposed to the lever associated with 1 pellet of food) across different values of the probe stimuli (a standardised scale, with *1.0* corresponding to the reference tone associated with 1 pellet of food, and *2.0* corresponding to the reference tone associated with 2 pellets of food), by *measurement phase / treatment / contingency* group. These predicted lines were generated from the model and estimates in Equation 0.34.

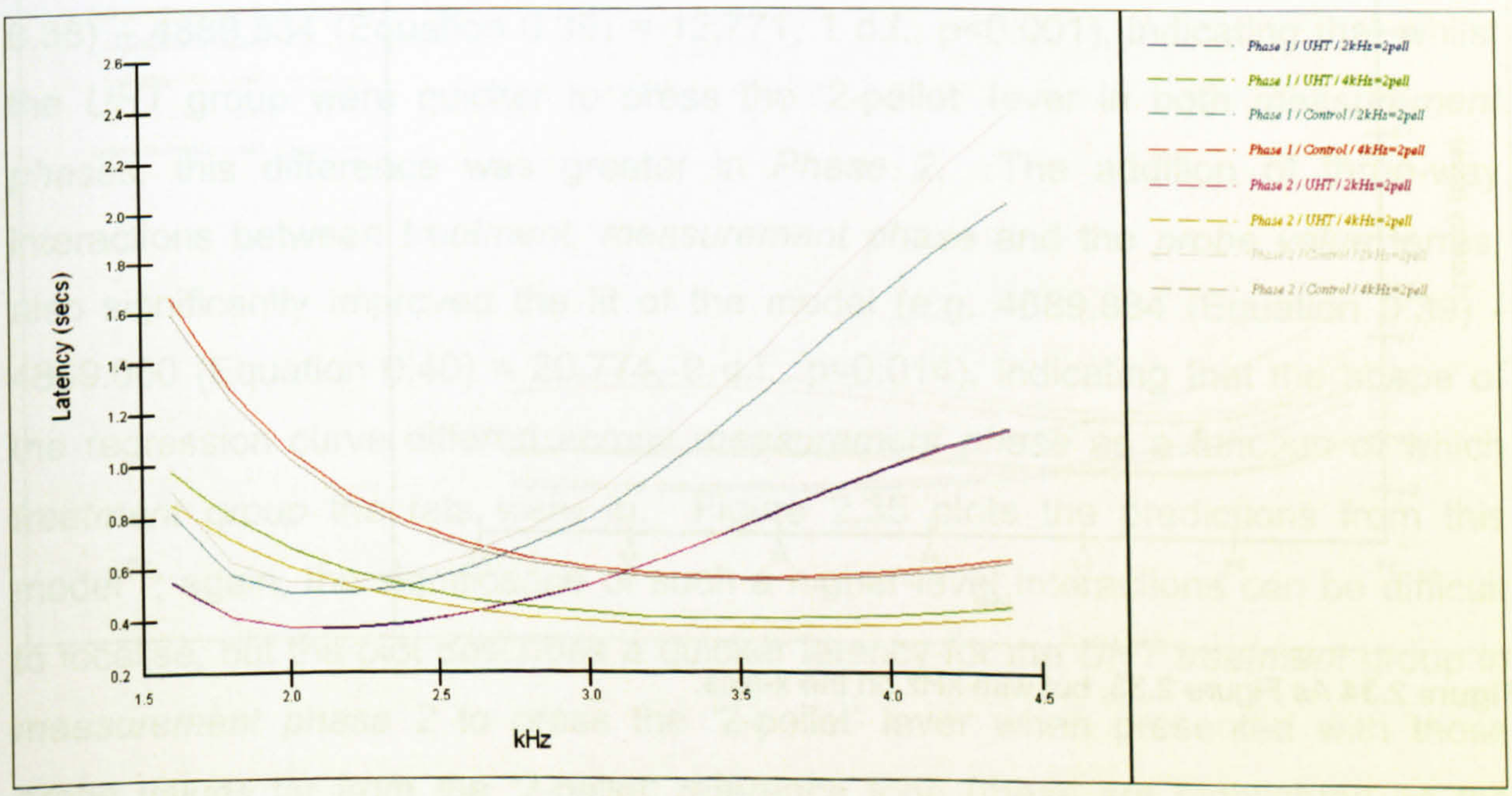

Figure 2.32 As Figure 2.31, but with kHz on the x-axis.

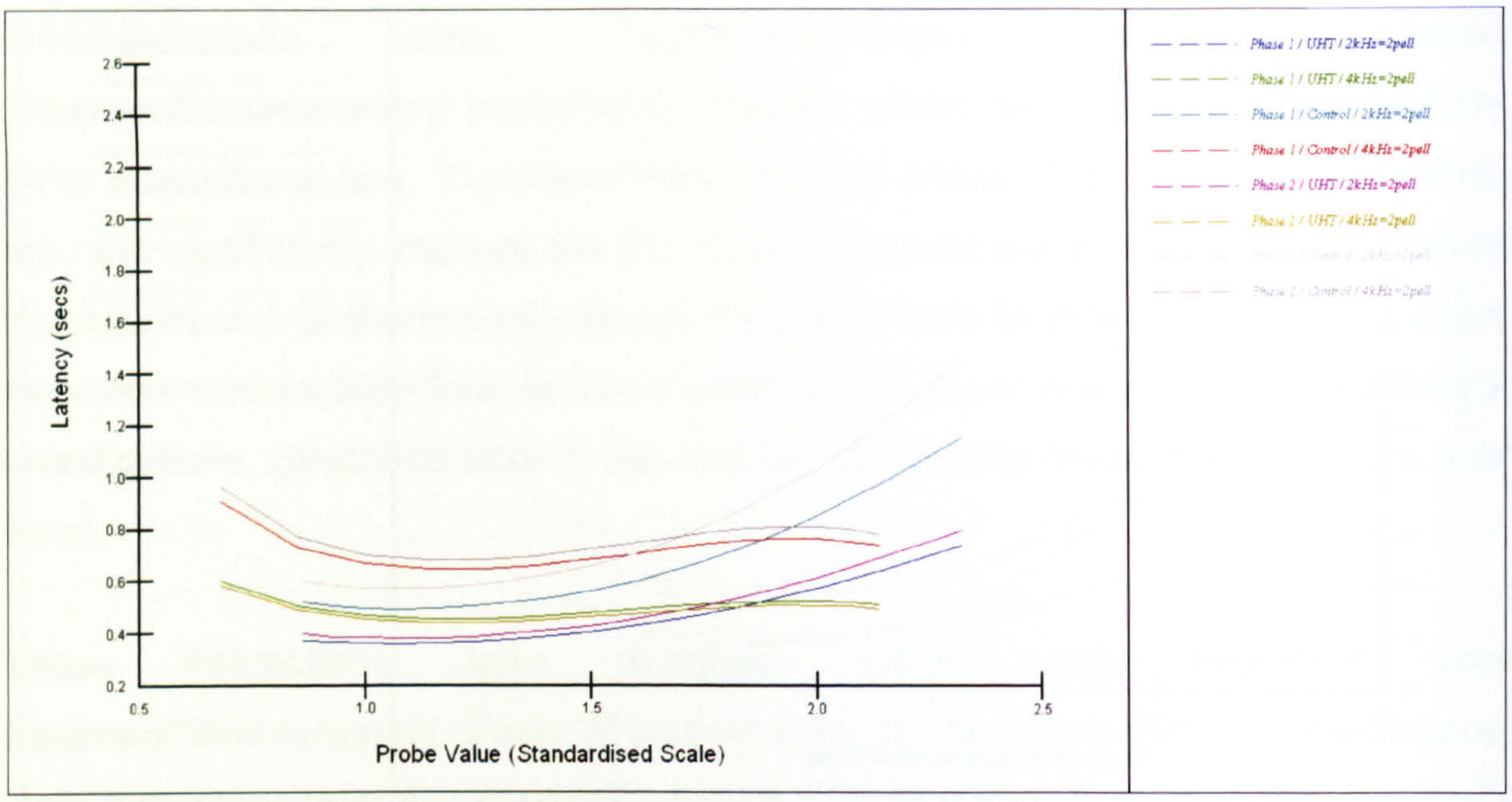

Figure 2.33 The latency (in seconds) to press the lever associated with 1 pellet of food (as opposed to the lever associated with 2 pellets of food) across different values of the probe stimuli (a standardised scale, with *1.0* corresponding to the reference tone associated with 1 pellet of food, and *2.0* corresponding to the reference tone associated with 2 pellets of food), by *measurement phase / treatment / contingency* group. These predicted lines were generated from the model and estimates in Equation 0.34.

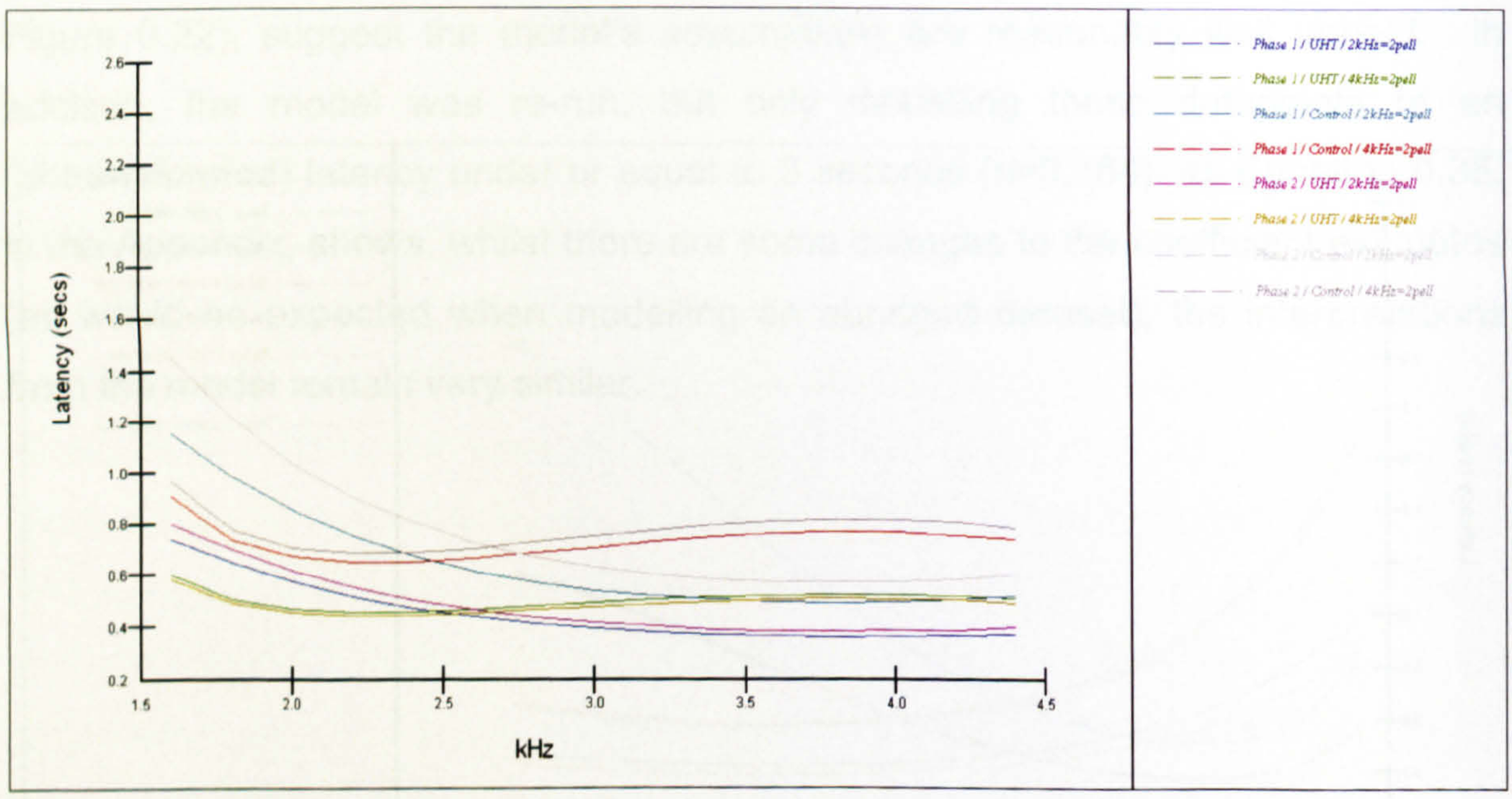

Figure 2.34 As Figure 2.33, but with kHz on the x-axis.

Finally, as in the multilevel analysis with *lever choice* as the response (y) variable (described above), the *Position of the 2-pellet lever* was added to the model, together with the *Room* in which the rats were tested, and the interaction between

the two terms, however, none of these modifications improved the fit of the model (e.g. 9772.200 - 9772.140 = 0.060, 1 d.f., p=0.807; Equation 0.36, in the Appendix).

As mentioned earlier, we produced models with quite a high degree of complexity when building our regression equations, above, and it was difficult to distill the precise meaning of some of the higher-level interactions we encountered. Therefore, we conducted two subsidiary analyses, each modelling the data from only one of the two levers (i.e. the '2-pellet' or '1-pellet' lever, respectively), thus simplifying the models. In these subsidiary analyses, we used very similar model-fitting procedures to those described above, so rather than introduce them again in detail, we present a more brief account here, with the main points of interest.

The analysis of the '2-pellet' lever data (n=4,594) revealed that adding *treatment*, as a main effect, significantly improved the fit of the model (4907.670 (Equation 0.37) – 4902.605 (Equation 0.38) = 5.065, 1 d.f., p=0.024), with the *UHT* group quicker than the *control* group to press the '2-pellet' lever. There was also a highly-significant *treatment*measurement phase* interaction (4902.605 (Equation 0.38) – 4889.834 (Equation 0.39) = 12.771, 1 d.f., p<0.001), indicating that whilst the *UHT* group were quicker to press the '2-pellet' lever in both *measurement phases,* this difference was greater in *Phase 2.* The addition of three-way interactions between *treatment, measurement phase* and the *probe value* terms, also significantly improved the fit of the model (e.g. 4889.834 (Equation 0.39) - 4869.060 (Equation 0.40) = 20.774, 9 d.f., p=0.014), indicating that the shape of the regression curve differed across *measurement phase* as a function of which *treatment group* the rats were in. Figure 2.35 plots the predictions from this model[76]; again, the significance of such a higher-level interactions can be difficult to localise, but the plot describes a quicker latency for the *UHT treatment* group in *measurement phase 2* to press the '2-pellet' lever when presented with those *probe values* far from the '2-pellet' reference tone (these are highlighted on the chart), whereas there is no such trend for the *control treatment* group (Figure 0.23

[76] We only present the plot with kHz on the x-axis (rather than the standardised probe value scale), as it is easier to distinguish each line in the chart.

to Figure 0.25, in the Appendix, plot the residuals from this equation; these suggest the model's assumptions were reasonably well upheld).

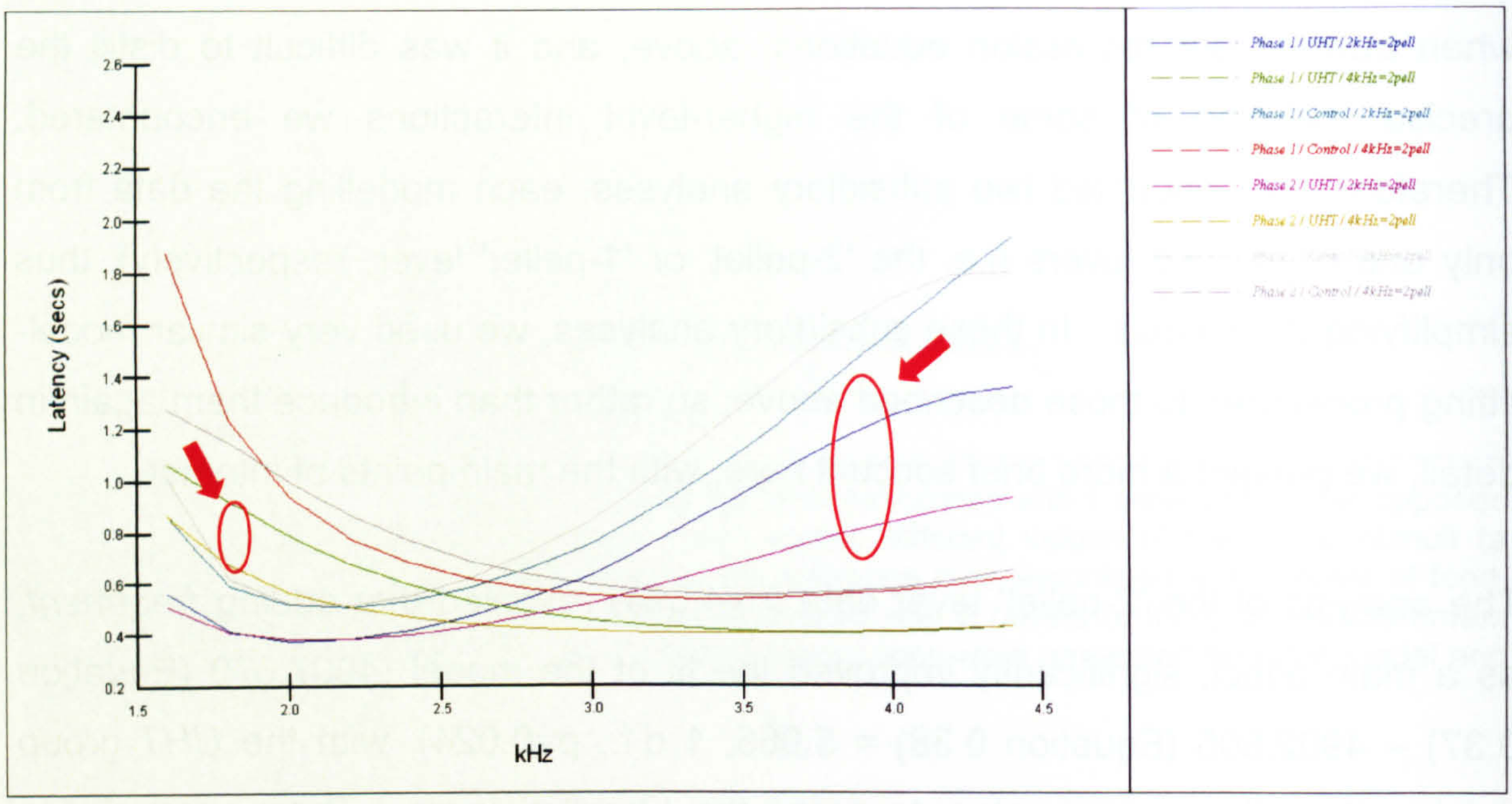

Figure 2.35 The latency (in seconds) to press the lever associated with 2 pellets of food, across kHz, by *measurement phase / treatment / contingency* group; the red arrows and loops refer to interactions covered in the main text. These predicted lines were generated from the model and estimates in Equation 0.40, which only modelled the '2-pellet' lever data.

The analysis of the '1-pellet' lever data found that the addition of *measurement phase* significantly improved the model's fit (4752.508 (Equation 0.41) - 4741.132 (Equation 0.42) = 11.376, 1 d.f., $p<0.001$), with presses on the '1-pellet' lever overall slower in *Phase 2,* compared to *Phase 1.* There was also a significant main effect of *treatment* (4726.413 (Equation 0.43) – 4719.563 (Equation 0.44) = 6.85, 1 d.f., $p=0.009$) with the *UHT* group pressing the lever more quickly, but, unlike the analysis of the '2-pellet' lever data, the interaction between *treatment* and *measurement phase* was not significant (4719.563 (Equation 0.44) - 4718.442 (Equation 0.45) = 1.121, 1 d.f., $p=0.290$), nor were three-way interactions between *treatment, measurement phase* and *probe value* (4738.659 (Equation 0.46) –

4728.050 (Equation 0.47) = 10.609, 7 d.f., p=0.157)[77]. Otherwise, the fit of the model was significantly improved by the addition of a two-way interaction between *measurement phase* and *contingency* (4741.132 (Equation 0.42) – 4726.413 (Equation 0.43) = 14.719, 1 d.f., p<0.001; indicating that in the second *measurement phase,* the *2kHz=2pell contingency* group were slower, and the *4kHz=2pell contingency* group were faster, than in the first *measurement phase*), and a three-way interaction between *measurement phase, contingency* and *treatment* (4719.562 (Equation 0.48) – 4701.103 (Equation 0.49) = 18.459, 3 d.f., p<0.001). Figure 2.36 plots the predictions from a model featuring this latter interaction, indicating that all the *treatment / contingency* subgroups are slower to press the '1-pellet' lever in the second *measurement phase,* except the *UHT / 4kHz=2pell* group, which are faster. Figure 0.27 and Figure 0.30, in Appendix, plot the residuals from this model, suggesting its assumptions are reasonably well met.

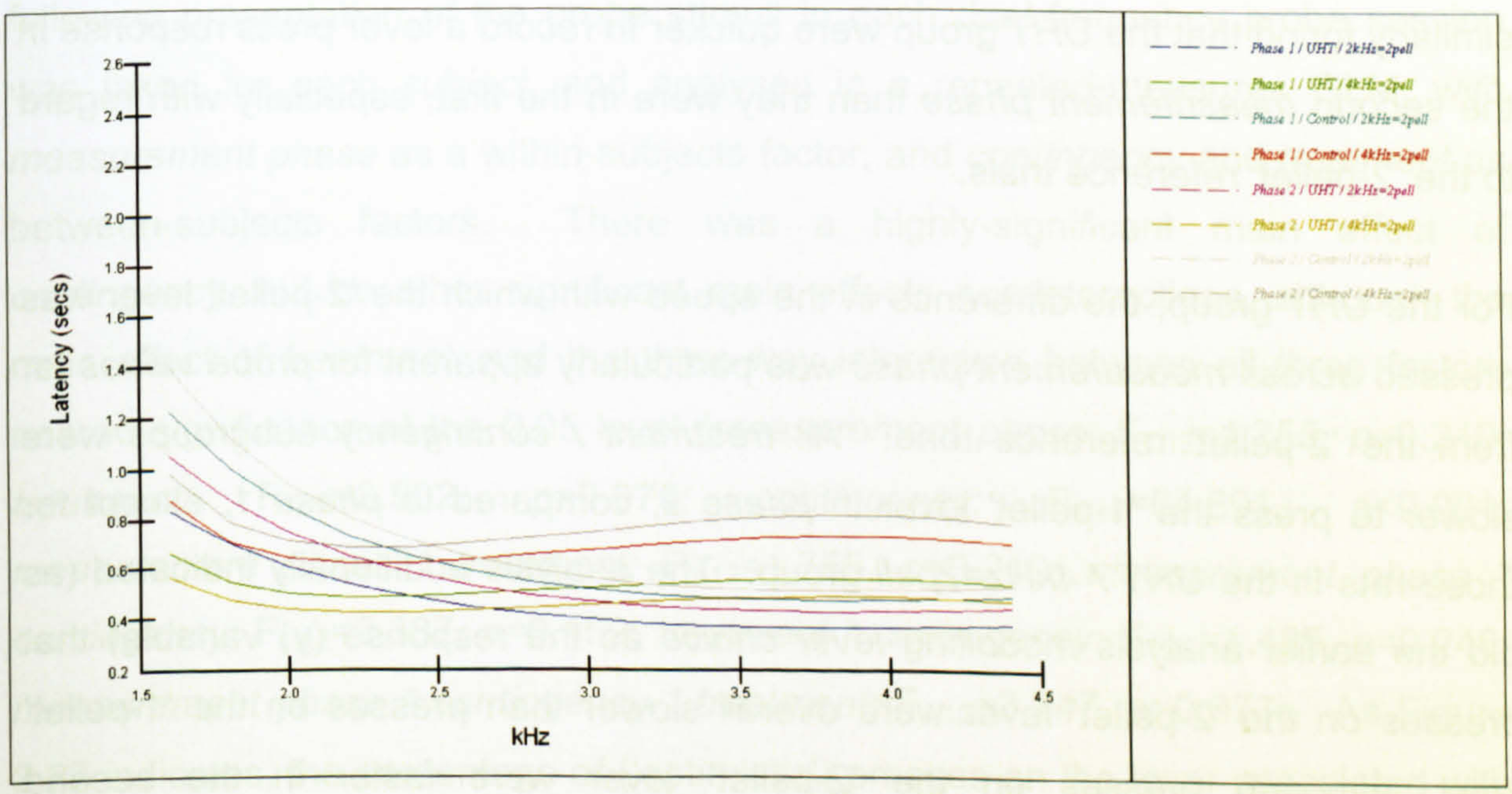

Figure 2.36 The latency (in seconds) to press the lever associated with 1 pellet of food, across kHz, by *measurement phase / treatment / contingency* group. These predicted lines were generated from the model and estimates in Equation 0.49, which only modelled the '1-pellet' lever data.

[77] N.B. the cubic terms were deleted from this model to allow it to converge.

Summary of the single-frequency probe latency analyses

The multi-level analysis conducted in MLwiN found that the *UHT* group were overall significantly quicker to record a lever press: both with regard to pressing *any* lever, and also with regard to pressing each individual lever (i.e. they were quicker to press the '2-pellet' lever, and also quicker to press the '1-pellet' lever). This difference between the two *treatment* groups was particularly pronounced in the second *measurement phase,* both with regard to the latency to press *any* lever (the *Control* group were slower in *phase 2* than they were in *phase 1,* whilst the *UHT* group were slightly faster in *phase 2* compared to *phase 1*), and specifically with regard to the latency to press the '2-pellet' lever (the *Control* group were slightly slower to press this lever in *phase 2* compared to *phase 1,* whilst the *UHT* group were considerably faster in *phase 2* than they were in *phase 1*). Incidentally, the analysis of the operant *training* data, described towards the start of this section, similarly found that the *UHT* group were quicker to record a lever press response in the second *measurement phase* than they were in the first, especially with regard to the '2-pellet' reference trials.

For the *UHT* group, the difference in the speed with which the '2-pellet' lever was pressed across *measurement phase* was particularly apparent for *probe values* far from the '2-pellet' reference tone. All *treatment / contingency* subgroups were slower to press the '1-pellet' lever in *phase 2,* compared to *phase 1,* except for those rats in the *UHT / 4kHz=2pell* group. The analysis additionally indicated (as did the earlier analysis modelling *lever choice* as the response (y) variable) that presses on the '2-pellet' lever were overall slower than presses on the '1-pellet' lever, although presses on the '2-pellet' lever were faster in the second *measurement phase* than they were in *phase 1,* and *vice versa* for presses on the '1-pellet' lever.

The analysis of the latency data conducted in MLwiN was generally in good agreement with an earlier, exploratory analysis using repeated-measures ANOVAS in SPSS (see Appendix C). The ANOVAS found non-significant trends for the *UHT* group to be overall faster to press *any* lever, and also to specifically press the '2-pellet' lever. In addition, each type of analysis found that the *2kHz=2pell*

contingency group were slower, and the *4kHz=2pell* group were faster, to press the '1-pellet' lever in the second *measurement phase* compared to the first (at least to a near-significant lever in the repeated-measures analyses). Finally, the plots of the data accompanying the repeated-measures ANOVAS suggested the difference between the two *treatment* groups in *latency*, at least with respect to pressing *any* lever, and with respect to pressing the '1-pellet' lever, was greater in the second *measurement phase* than in the first, although none of the interactions between *treatment* and *measurement phase* were significant in those analyses.

Dual-frequency probe trials

Lever choice

The mean percentage of presses on the lever associated with 2 pellets of food following presentation of the probe stimuli in each dual-frequency probe session was taken for each *subject*, and analysed in a repeated-measures GLM, with *measurement phase* as a within-subjects factor, and *contingency* and *treatment* as between-subjects factors. There was a highly-significant main effect of *contingency*, but no other significant main effects nor interactions, although the main effect of *treatment*, and the three-way interaction between all three factors neared significance at the 0.05 level (*measurement phase:* $F_{1,12}$=1.755, p=0.210; *treatment:* $F_{1,12}$=3.802, p=0.075; *contingency:* $F_{1,12}$=64.691, p<0.001; *measurement phase * treatment:* $F_{1,12}$=1.755, p=0.210; *measurement phase * contingency:* $F_{1,12}$=2.337, p=0.152; *treatment * contingency:* $F_{1,12}$=1.485, p=0.246; *measurement phase * contingency * treatment:* $F_{1,12}$=3.747, p=0.077). As Figure 2.37 indicates, the percentage of ('optimistic') presses on the lever associated with 2 pellets of food is consistently greater for the *4kHz=2pell contingency* group. The responding of the various *treatment / contingency* subgroups is fairly consistent across *measurement phase*, except for the *Control / 2kHz=2pell* group which has a 'pessimistic' downshift in responding.

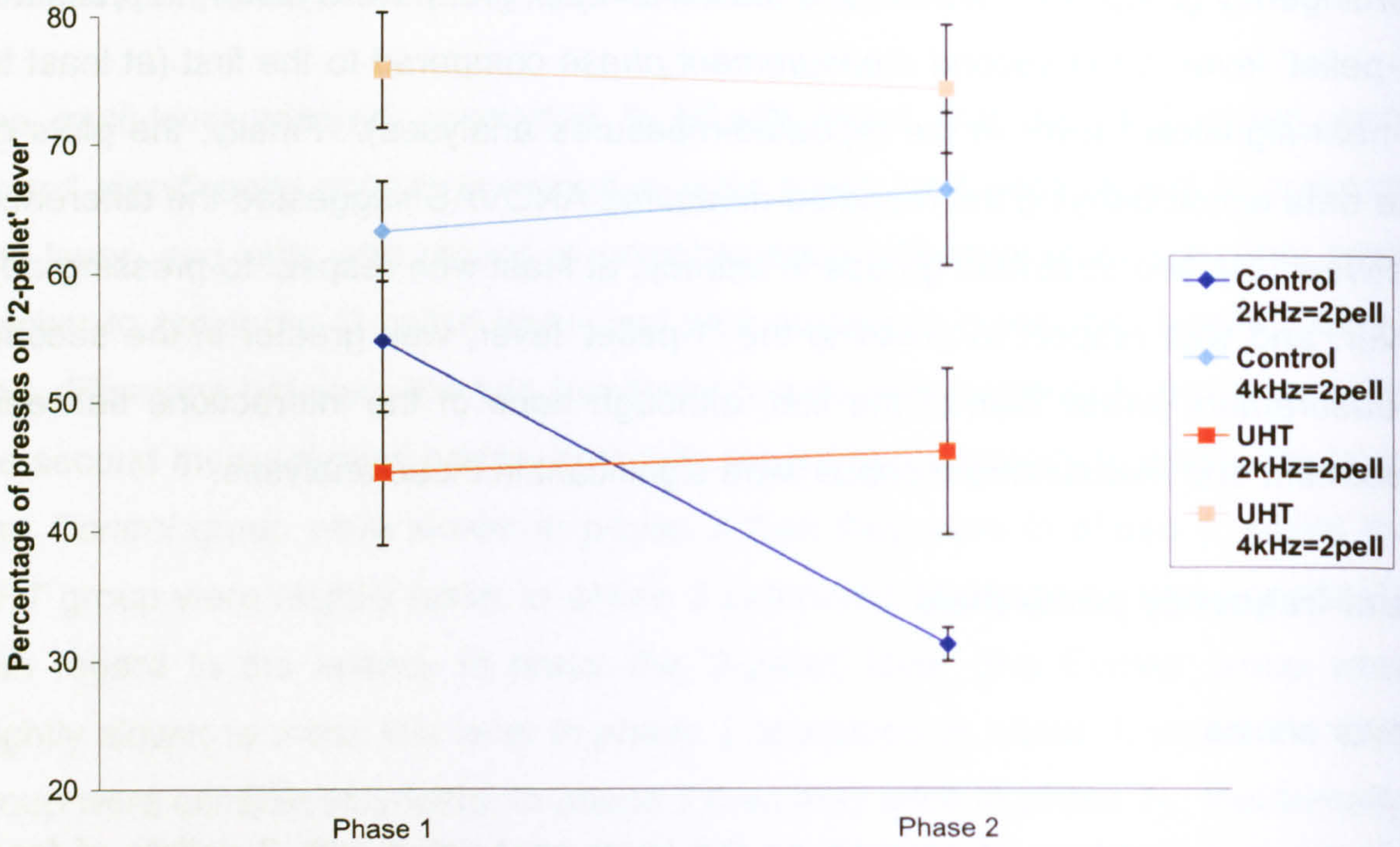

Figure 2.37 The mean percentage of presses on the lever associated with 2 pellets of food following presentation of the probe tones in the dual-frequency probe sessions for each of the *Treatment / Contingency* groups, across *Measurement phase* (+/- 1SEM).

Latency

The mean *latency* to press either lever (i.e. regardless of that lever's identity) in the probe trials was taken for each *subject* for each dual-frequency test session, and analysed (following a negative inverse-transformation) in a repeated-measures GLM of the same design as that employed above. *Treatment,* as a main effect, was significant, as was the two-way interactions between *measurement phase* and *contingency*; all other main effects and interactions were non-significant (*measurement phase:* $F_{1,12}$=1.268, p=0.282; *treatment:* $F_{1,12}$=7.605, p=0.017; *contingency:* $F_{1,12}$=2.546, p=0.137; *measurement phase * treatment:* $F_{1,12}$=0.006, p=0.942; *measurement phase * contingency:* $F_{1,12}$=9.675, p=0.009; *treatment * contingency:* $F_{1,12}$=1.644, p=0.224; *measurement phase * contingency * treatment:* $F_{1,12}$=3.645, p=0.080). Figure 2.38 plots the mean *latency* across *measurement phase* for each of the *treatment* groups, indicating that the UHT group are consistently quicker to make a lever-press response. Figure 2.39, which plots the same data, but summarised by *contingency* group, indicates that whilst each of the *contingency* groups had similar latencies in *phase 1,* the *4kHz=2pell* group are

quicker to record their lever responses in *phase 2*, whilst the *2kHz=2pell* group are slightly slower than they were in *phase 1*.

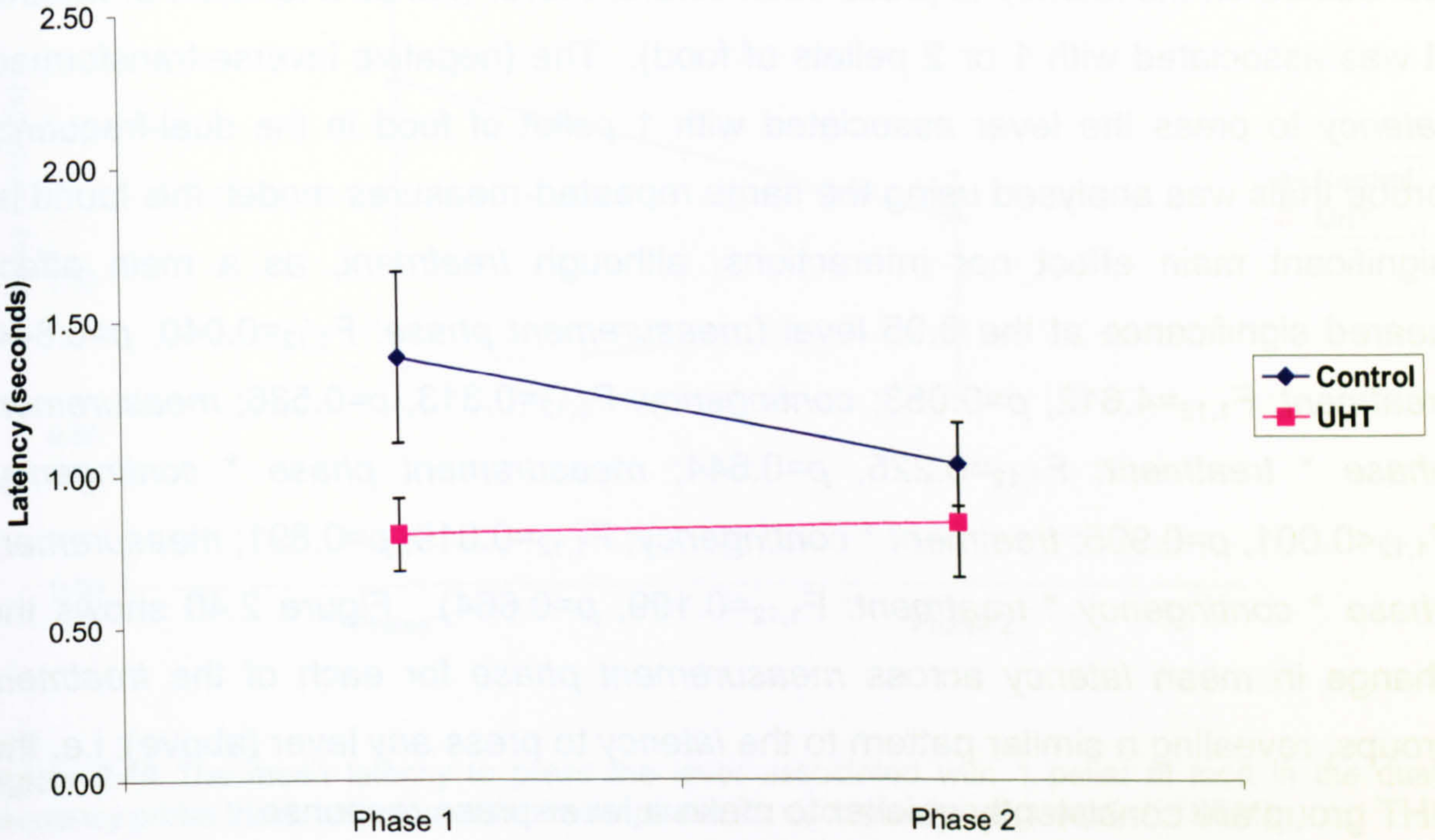

Figure 2.38 The mean latency to press either lever in the dual-frequency probe trials, across *measurement phase* for each *treatment* group (+/- 1SEM).

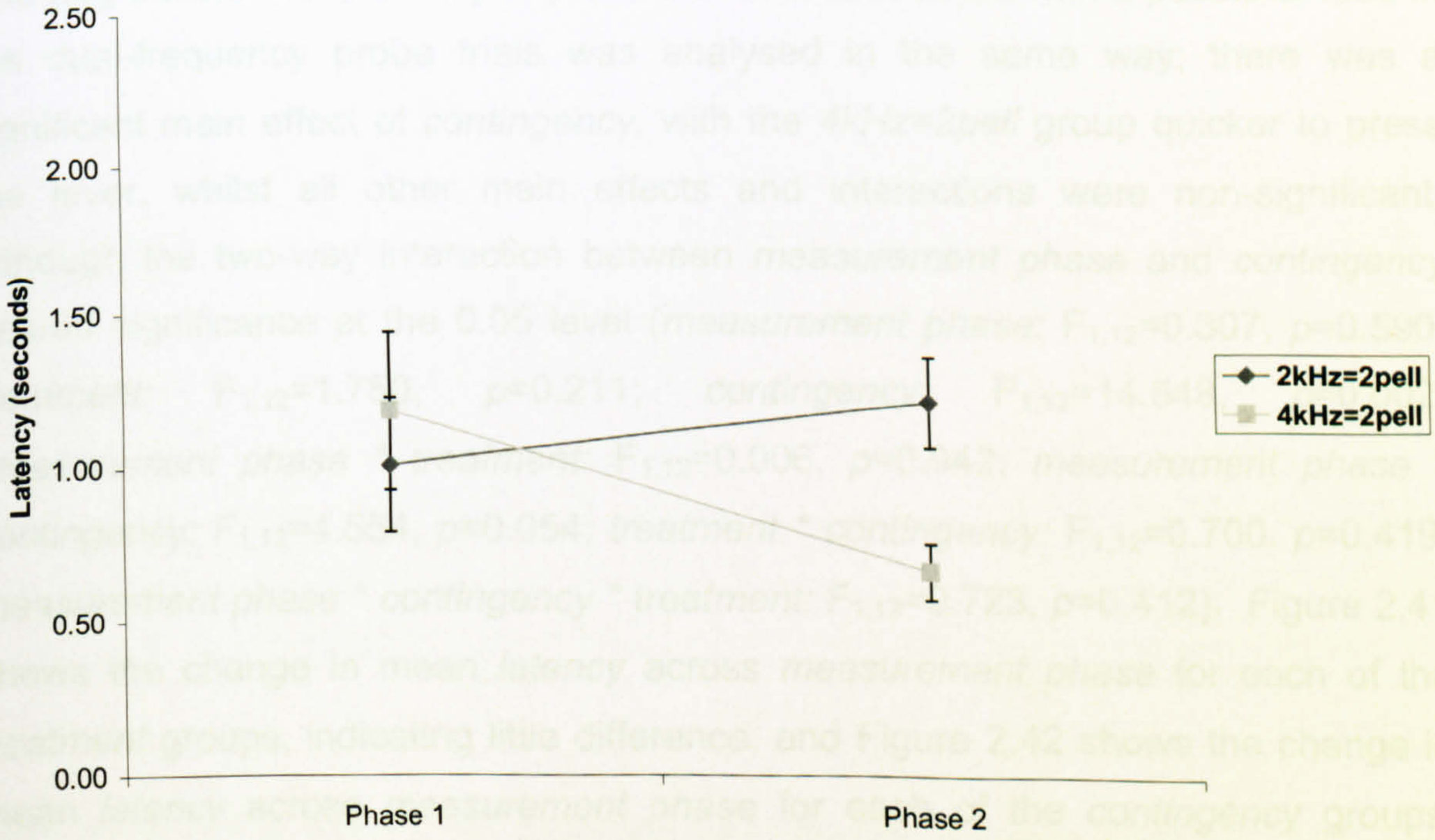

Figure 2.39 The mean latency to press either lever in the dual-frequency probe trials, across *measurement phase* for each *contingency* group (+/- 1SEM).

As in the analysis for the single-frequency probe tests, additional analyses were conducted on the *latency* to press each *different* lever (i.e. as a function of whether it was associated with 1 or 2 pellets of food). The (negative inverse-transformed) *latency* to press the lever associated with 1 pellet of food in the dual-frequency probe trials was analysed using the same repeated-measures model; this found no significant main effect nor interactions, although *treatment,* as a main effect, neared significance at the 0.05 level (*measurement phase:* $F_{1,12}$=0.040, p=0.844; *treatment:* $F_{1,12}$=4.612, p=0.053; *contingency:* $F_{1,12}$=0.313, p=0.586; *measurement phase * treatment:* $F_{1,12}$=0.225, p=0.644; *measurement phase * contingency:* $F_{1,12}$<0.001, p=0.995; *treatment * contingency:* $F_{1,12}$=0.019, p=0.891; *measurement phase * contingency * treatment:* $F_{1,12}$=0.199, p=0.664). Figure 2.40 shows the change in mean *latency* across *measurement phase* for each of the *treatment* groups, revealing a similar pattern to the *latency* to press *any* lever (above): i.e. the UHT group are consistently quicker to make a lever-press response.

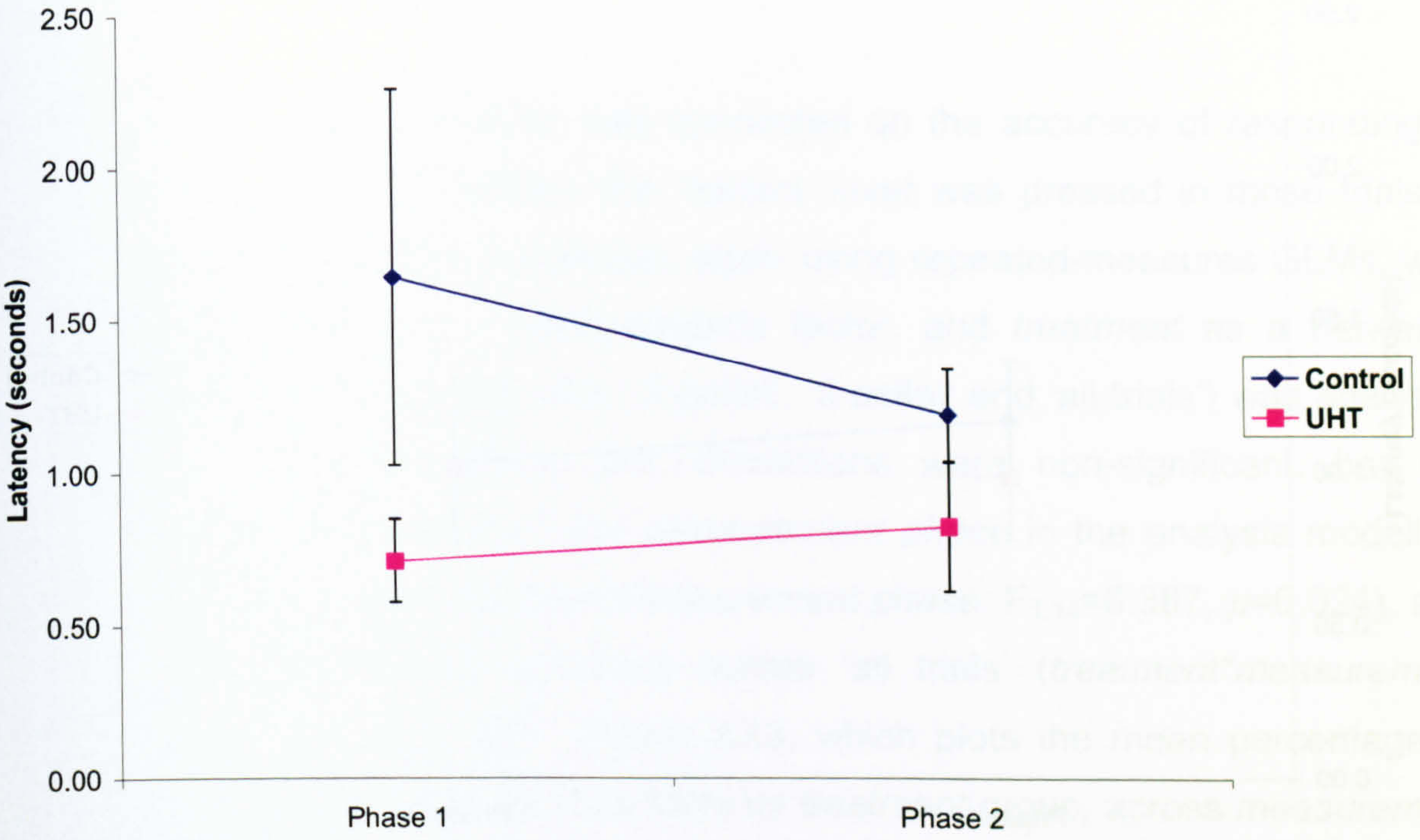

Figure 2.40 The mean latency to press the lever associated with 1 pellet of food in the dual-frequency probe trials, across *measurement phase* for each *treatment* group (+/- 1SEM).

The (log-transformed) *latency* to press the lever associated with 2 pellets of food in the dual-frequency probe trials was analysed in the same way; there was a significant main effect of *contingency,* with the *4kHz=2pell* group quicker to press the lever, whilst all other main effects and interactions were non-significant, although the two-way interaction between *measurement phase* and *contingency* neared significance at the 0.05 level (*measurement phase:* $F_{1,12}$=0.307, p=0.590; *treatment:* $F_{1,12}$=1.750, p=0.211; *contingency:* $F_{1,12}$=14.548, p=0.002; *measurement phase * treatment:* $F_{1,12}$=0.006, p=0.942; *measurement phase * contingency:* $F_{1,12}$=4.554, p=0.054; *treatment * contingency:* $F_{1,12}$=0.700, p=0.419; *measurement phase * contingency * treatment:* $F_{1,12}$=0.723, p=0.412). Figure 2.41 shows the change in mean *latency* across *measurement phase* for each of the *treatment* groups, indicating little difference, and Figure 2.42 shows the change in mean *latency* across *measurement phase* for each of the *contingency* groups, indicating that the *4kHz=2pell* group are consistently quicker to press the '2-pellet' lever in the dual-frequency probe trials.

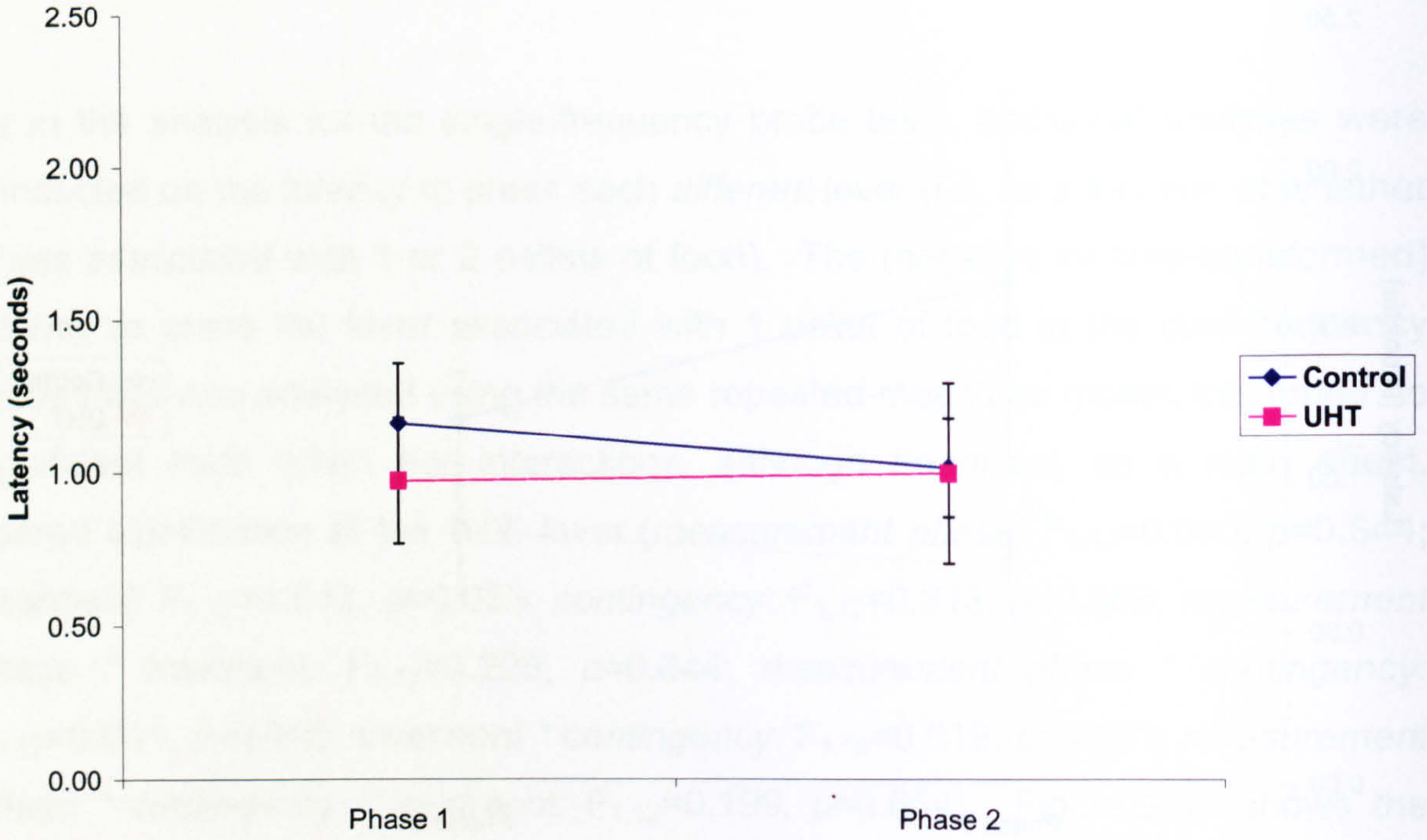

Figure 2.41 The mean latency to press the lever associated with 2 pellets of food in the dual-frequency probe trials, across *measurement phase* for each *treatment* group (+/- 1SEM).

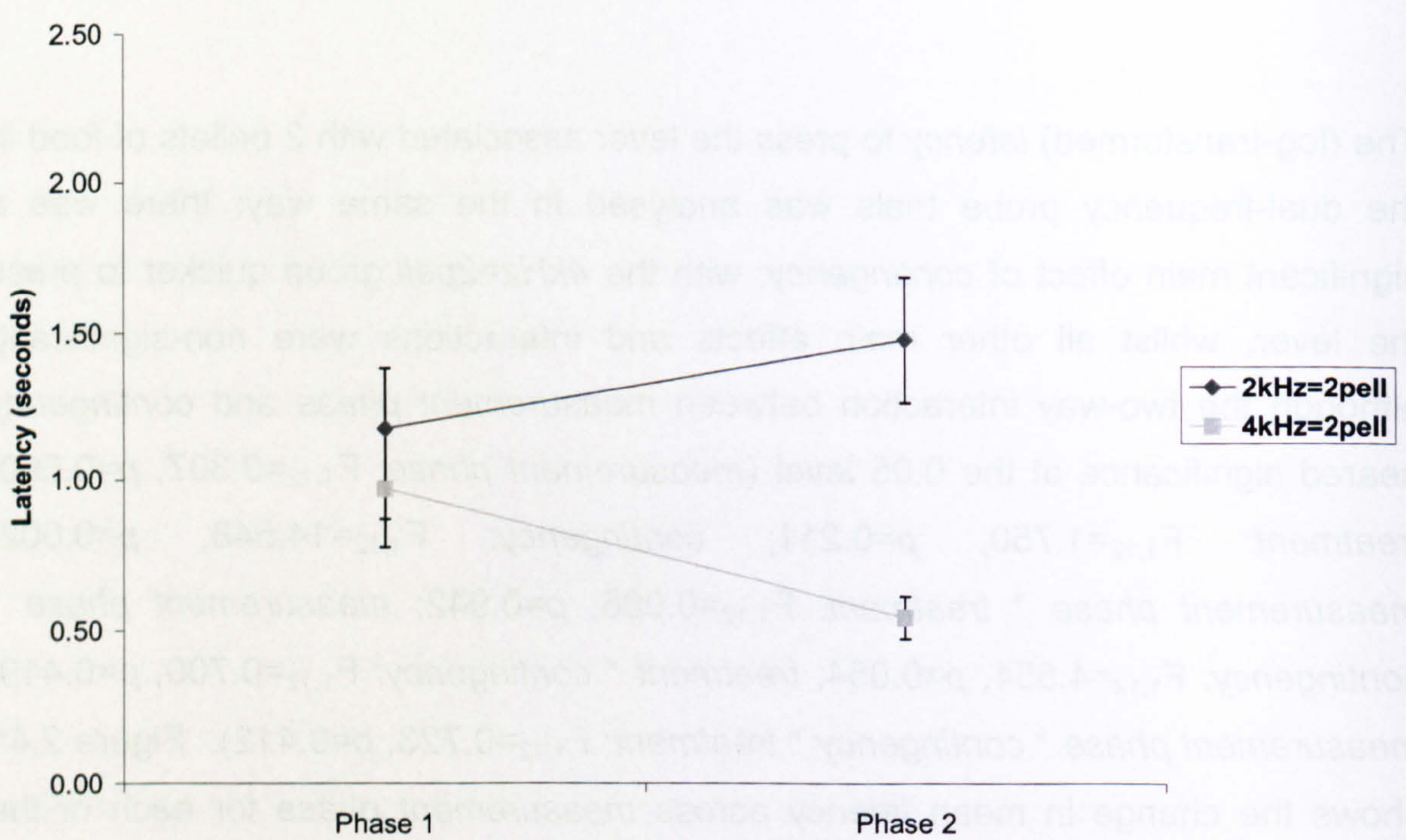

Figure 2.42 The mean latency to press the lever associated with 2 pellets of food in the dual-frequency probe trials, across *measurement phase* for each *contingency* group (+/- 1SEM).

Accuracy in reference trials

Finally, an exploratory analysis was conducted on the accuracy of responding in the reference trials (i.e. whether the 'correct' lever was pressed in those trials in which 2kHz *or* 4kHz were presented), again using repeated-measures GLMs, with *measurement phase* as a within-subjects factor, and *treatment* as a between-subjects factor; each trial type (i.e. '1-pellet', '2-pellet' and 'all trials') was analysed separately. All main effects and interactions were non-significant, bar the interaction between *treatment* and *measurement phase* in the analysis modelling '2-pellet' trial accuracy (*treatment*measurement phase:* $F_{1,14}$=6.387, p=0.024), and in the analysis modelling accuracy across 'all trials' (*treatment*measurement phase:* $F_{1,14}$=5.587, p=0.033). Figure 2.43, which plots the mean percentage of 'correct' responses in the reference trials by *treatment* group, across *measurement phase,* indicates that the *UHT* group were more accurate in all trials in the second measurement phase, compared to the first, whereas the *Control* group were less accurate in the second *measurement phase*, compared to the first, across *all* trials, and also specifically in the '2-pellet' trials.

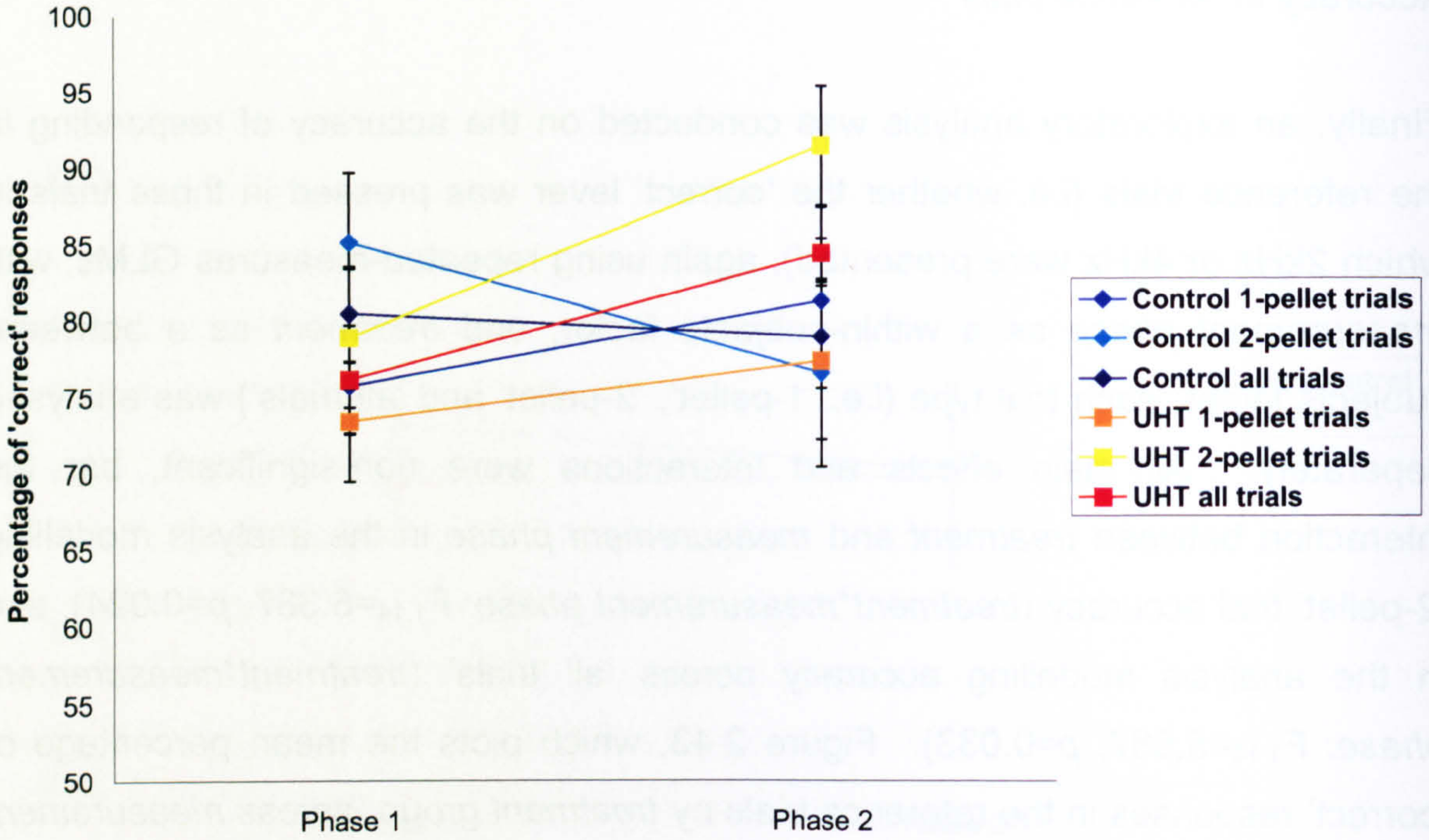

Figure 2.43 The mean percentage of 'correct' responses in the reference trials of the dual-frequency probe sessions, in each of the two *measurement phases*, by *treatment.* The data are summarised by trial type (+/- 1SEM).

Summary of dual-frequency probe analyses

As in the analyses of the single-frequency probe data, the analysis of *lever choice* in the dual-frequency probe sessions found a (near-significant) tendency for the *2kHz=2pell contingency* rats in the *Control* group to respond more 'pessimistically' in *measurement phase 2,* compared to *phase 1*, indicating that the change in their response pattern across *measurement phase* remained relatively consistent across the different types of probe-testing (i.e. across both single-frequency and dual-frequency probe test sessions). More generally, the *2kHz=2pell contingency* group were overall more 'pessimistic' when responding to the dual-frequency probe trials, as found in the analyses of the *lever choice* data from the single-frequency probe sessions.

Again, as in the analyses of the single-frequency probe data, the *UHT* group were significantly faster to record a lever press response in the dual-frequency probe trials, at least when all lever presses were grouped together (and also to a near-

significant level with regard to only those presses on the '1-pellet' lever). Otherwise, the *4kHz=2pell contingency* group were overall faster to press the '2-pellet' lever. In addition, they were also faster to press *any* lever in the second *measurement phase,* compared to the first, whereas the opposite was true of the *2kHz=2pell* group (this interaction also neared-significance for the latency to specifically press the '2-pellet' lever); again, the significance of this interaction is in keeping with that found in the analyses of latency pertaining to the single-frequency probe test sessions.

Finally, the analysis of response accuracy in the reference trials indicated that the *UHT* group were overall more accurate in the second *measurement phase,* compared to the first, whereas the opposite was true for the *Control* group; this significant interaction was also found when solely analysing the '2-pellet' trials.

Concurrent tests

Time taken to eat 50 pellets of food

The time taken, in seconds, to eat 50 pellets of food was negative inverse-transformed, and submitted to a repeated-measures GLM, with *measurement phase* as a within-subject factor (with two levels: *pre-treatment* and *post-treatment*), and *treatment* as a between-subject factor. *Measurement phase* had a significant main effect, with a shorter latency to eat the 50 pellets *post-treatment,* whilst *treatment,* as a main effect, and the interaction between the two terms, were both non-significant (*measurement phase:* $F_{1,12}$=6.289, p=0.025; *treatment:* $F_{1,12}$=1.455, p=0.248; *measurement phase* * *treatment:* $F_{1,12}$=0.035, p=0.854). Figure 2.44 shows the mean latencies for each *treatment* group across *measurement phase*: both *treatment* groups eat the 50 pellets of food more quickly when tested the second time (i.e. *post-treatment*).

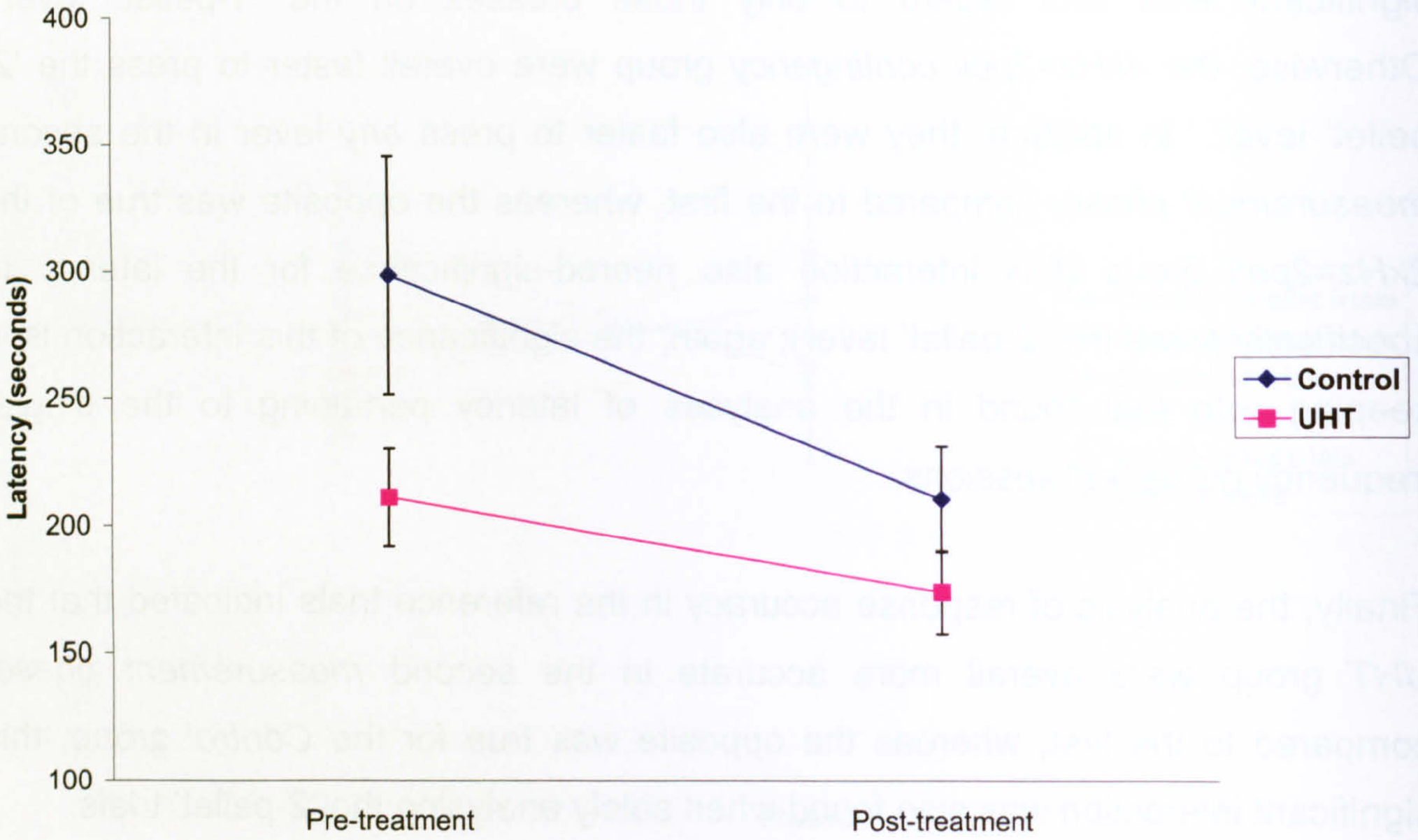

Figure 2.44 The mean latency (in seconds) to eat 50 pellets of freely-available food, by *treatment* group across *measurement phase* (+/- 1SEM).

Lever-based progressive ratio test with food reinforcement

All test sessions terminated before 60 minutes (i.e. the maximum session length) had elapsed. The total number of times the lever was pressed during the progressive ratio test session was taken for each *subject*, and submitted to an independent-samples *t*-test, which found no significant effect of *treatment* (t_{13}=0.289, p=0.778).[78] Figure 2.45 shows the mean number of lever presses for each *treatment* group.

[78] As mentioned in the Method section, one *subject* in the *control* group was excluded from this test due to ill-health on the test day.

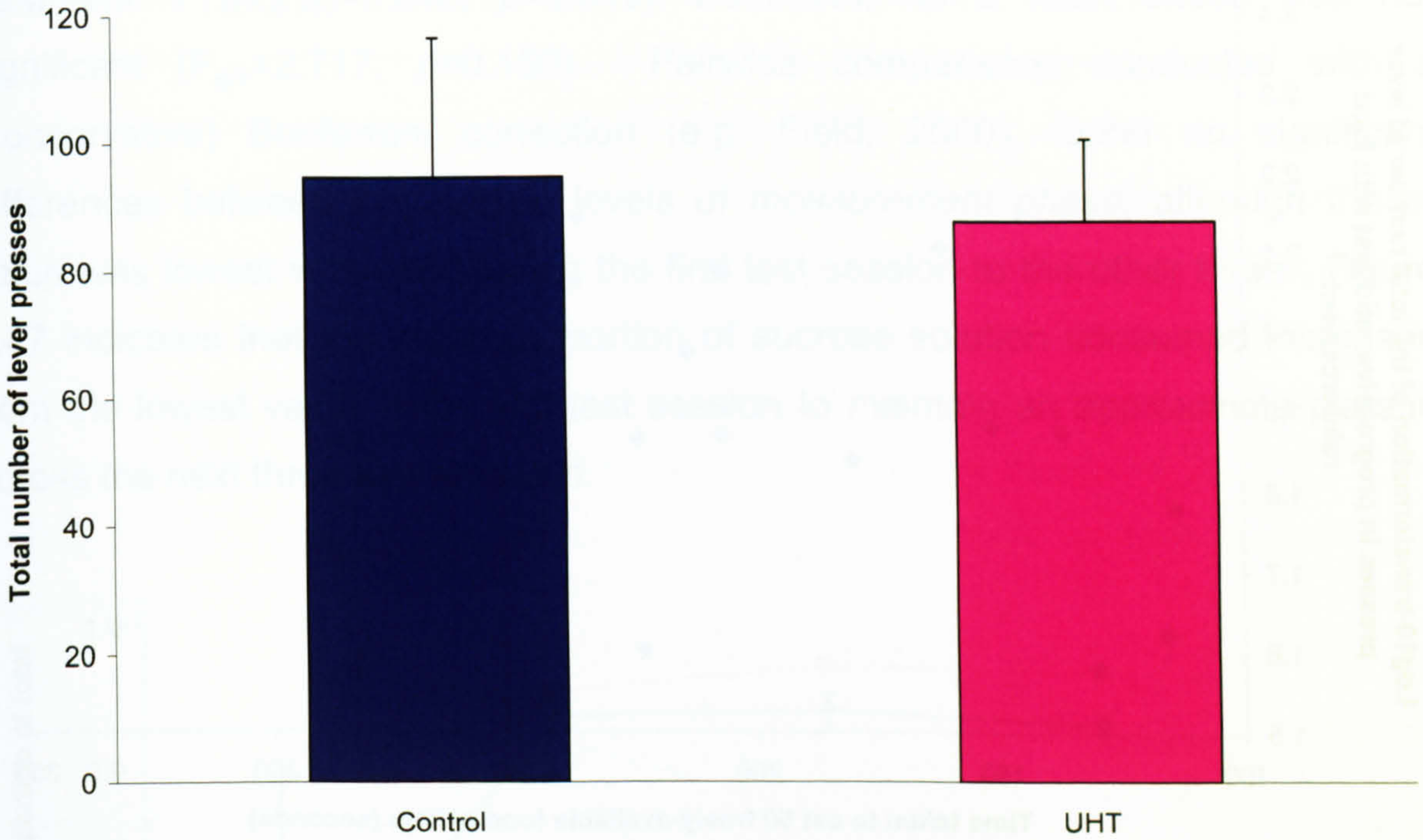

Figure 2.45 The mean number of times the lever was pressed in the progressive ratio test session by *treatment* group (error bar = 1SEM).

Tests of 'food motivation': concurrent validity

The latency to eat 50 pellets of food *post-treatment* was correlated with the (log-transformed) total number of lever presses made in the progressive ratio test session (which was only conducted *post-treatment*). This correlation was highly-significant (r=0.670, d.f.=13, p=0.006), indicating that the rats pressing the lever more during the progressive ratio session also took longer to eat 50 pellets of freely-available food. Figure 2.46 plots this relationship, and indicates the presence of an outlying datapoint with the highest value on each axis, which is some distance from the other datapoints; an additional analysis which excluded this point found the correlation neared significance at the 0.05 level (r=0.526, d.f.=12, p=0.053). This relationship is contrary to that hypothesised by good concurrent validity.

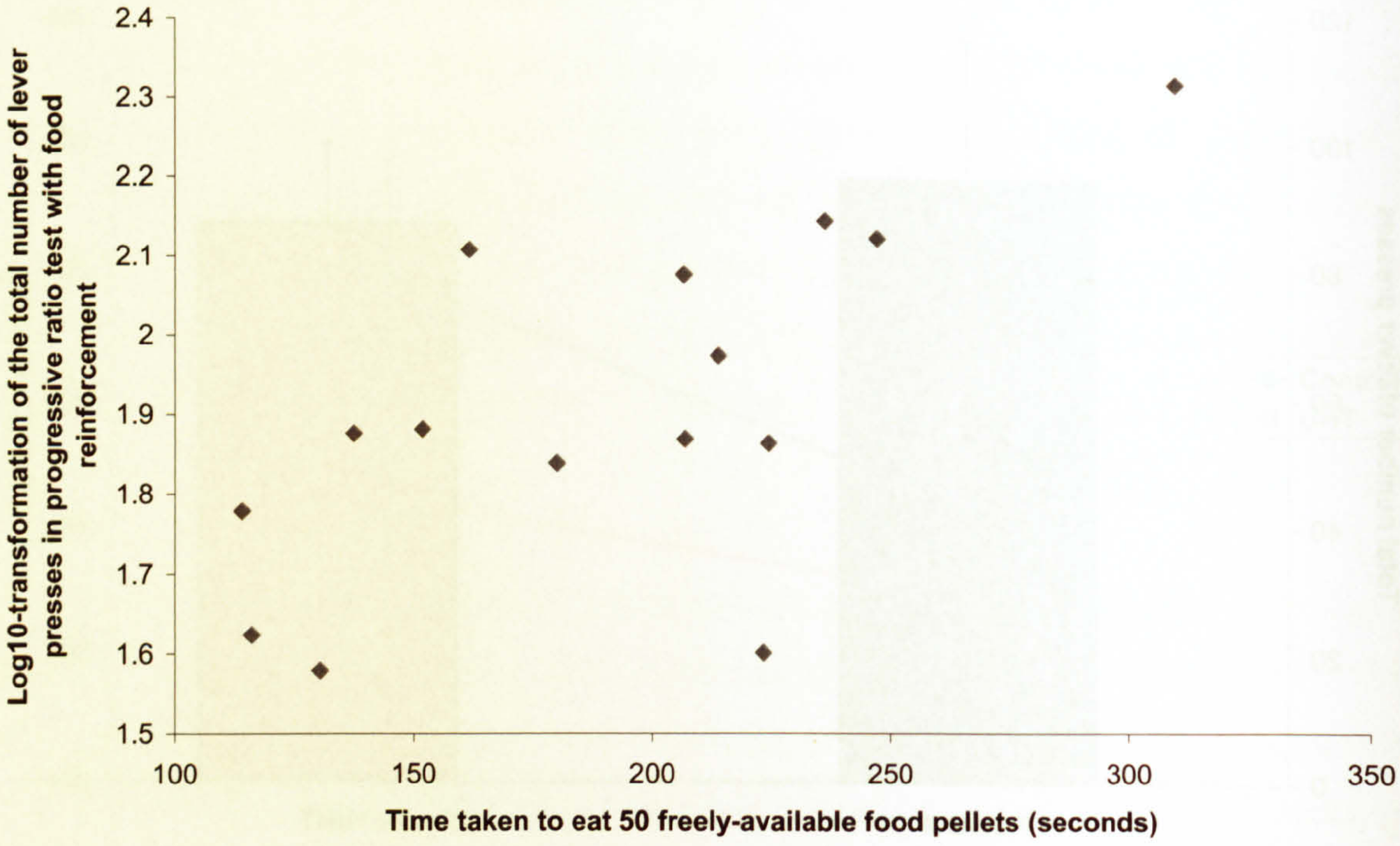

Figure 2.46 The (log-transformed) total number of times the lever was pressed in a progressive ratio test with food reinforcement against the time taken to eat 50 freely-available food pellets (in seconds).

Sucrose consumption: homecage-based test

The amount of sucrose solution consumed as an (arcsin-square-root-transformed) proportion of total fluid (i.e. water and sucrose solution) consumed in each homecage was submitted to a repeated-measures GLM, with *measurement phase* as a within-subjects factor (with four levels: a test session conducted *pre-treatment,* and three test sessions conducted *post-treatment*), and *treatment* as a between-subjects factor[79]. Inspection of the adjusted univariate output, and multivariate output, revealed the interaction between *measurement phase* and *treatment* was non-significant, whilst *measurement phase* was non-significant in the multivariate output, and bordered significance at the 0.05 level in the adjusted univariate output (e.g. the Greenhouse-Geisser univariate adjustment yielded *measurement phase:* $F_{1.055,6.333}=5.767$, $p=0.050$; and *measurement phase* *

[79] The residuals from the *pre-treatment measurement phase* failed the formal test of normality (e.g. Kolmogorov-Smirnov: p=0.049), and an inspection of the residual plots suggested this was due to a low-value outlier. None of the transformations we explored remedied this without compromising assumptions elsewhere, and the transformation we present in the text represented the best compromise.

treatment: $F_{1.055,6.333}$=0.048, p=0.846); *treatment*, as a main effect, was not significant ($F_{1,6}$=2.717, p=0.150). Pairwise comparisons conducted with a (conservative) Bonferroni correction (e.g. Field, 2000), found no significant differences between the various levels of *measurement phase,* although the *p*-value was lowest when comparing the first test session to the other three. Figure 2.47 indicates that the mean proportion of sucrose solution consumed increased from the lowest value in the first test session to maintain an approximate plateau across the next three test sessions.

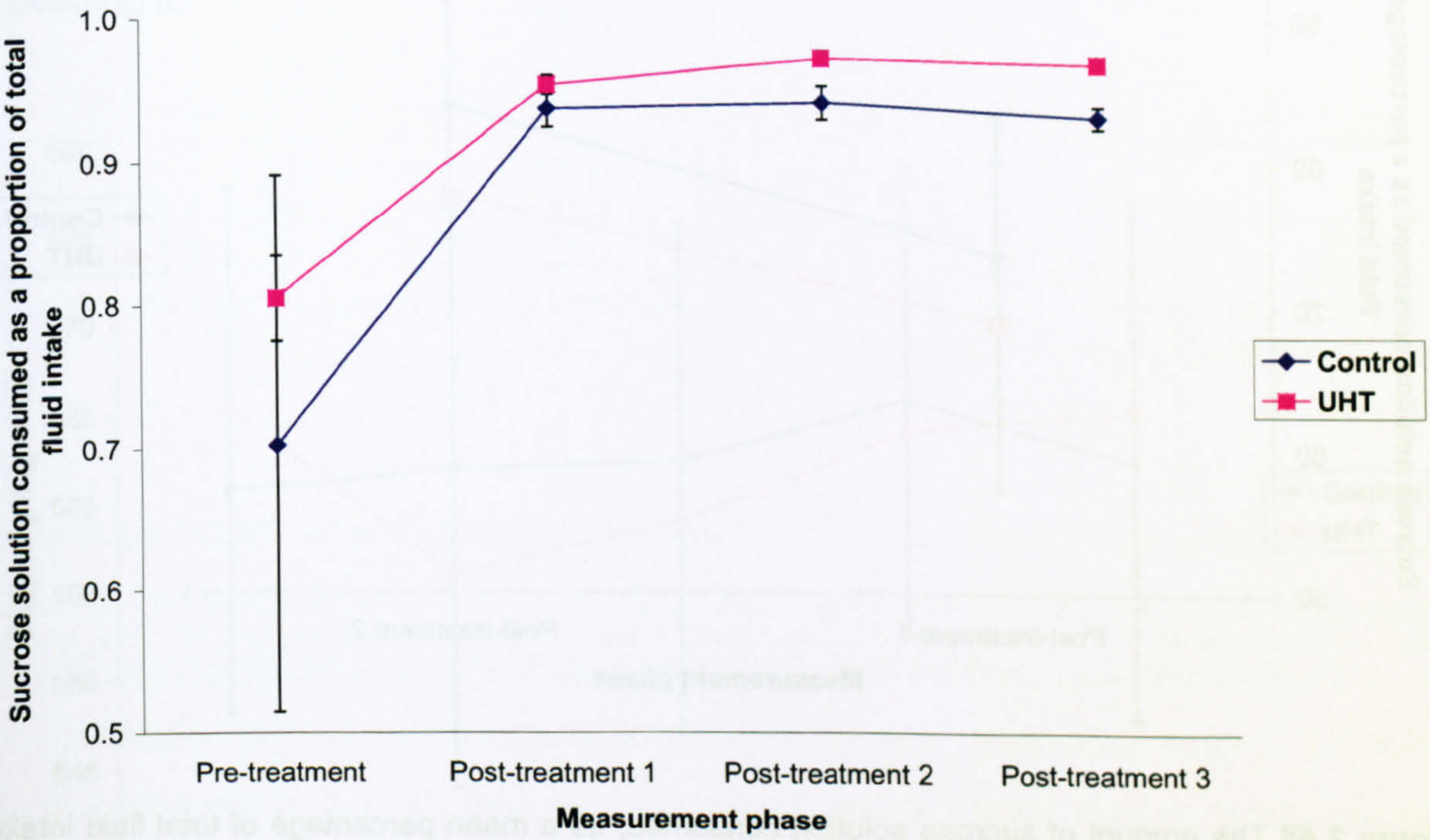

Figure 2.47 The mean amount of sucrose solution consumed as a proportion of total fluid intake (i.e. both water and sucrose solution) in a homecage-based sucrose preference test, by *treatment* across *measurement phase* (+/- 1SEM; since we used an arsin-square-root-transformation, which requires that the data be expressed as a proportion, we present proportions, rather than percentages, on the y-axis).

Sucrose consumption: individual-based test

The amount of sucrose solution consumed as a percentage of total fluid (i.e. water and sucrose solution) intake was reflected[80], then log-transformed prior to analysis

[80] i.e. adding one to the highest value in the dataset, then subtracting each value from this constant; if the data is subsequently log-transformed, for example, such 'reflection' can be a good solution for normalising data with a strong negative skew (e.g. Quinn & Keough, 2002).

in a repeated-measures GLM with *measurement phase* as a within-subjects factor (with two levels: *Post-treatment 1* and *Post-treatment 2*) and *treatment* as a between-subjects factor; none of the main effects nor interactions were significant (*measurement phase:* $F_{1,14}$=1.051, p=0.323; *treatment:* $F_{1,14}$=0.046, p=0.833; *measurement phase * treatment*: $F_{1,14}$=0.028, p=0.869). Figure 2.48 plots the mean percentages for each *treatment* group across *measurement phase.*

Figure 2.48 The amount of sucrose solution consumed, as a mean percentage of total fluid intake (i.e. water and sucrose solution) by *treatment* group across *measurement phase* (+/- 1SEM).

Bodyweights

The rats' bodyweight, in grams, was analysed in a repeated-measures GLM, with *measurement phase* as a within-subject factor (with five levels: *pre-treatment* and four occasions their bodyweight was measured *post-treatment*), and *treatment* as a between-subjects factor. Inspection of the multivariate and adjusted univariate output found that *measurement phase* was significant at the 0.05 level in both, whilst the interaction between *measurement phase* and *treatment* was significant only in the multivariate output (e.g. the Greenhouse-Geisser univariate adjustment yielded *measurement phase:* $F_{2.026,28.367}$=3.398, p=0.047; and *measurement phase * treatment:* $F_{2.026,28.367}$=2.807, p=0.077; the multivariate output yielded

measurement phase: $F_{4,11}$=10.313, *p*=0.001; and *measurement phase * treatment:* $F_{4,11}$=3.948, *p*=0.032); *treatment*, as a main effect, was not significant ($F_{1,14}$=0.002, *p*=0.967).

Figure 2.49 plots the mean bodyweight for each *treatment* group across *measurement phase*[81]. The Bonferroni-corrected pairwise comparisons between each day (i.e. the different levels of *measurement phase*), detailed in Table 2.19, indicated there were significant differences between *post-treatment 3*, in which the overall mean was highest, and the two *measurement phases* immediately preceding it.

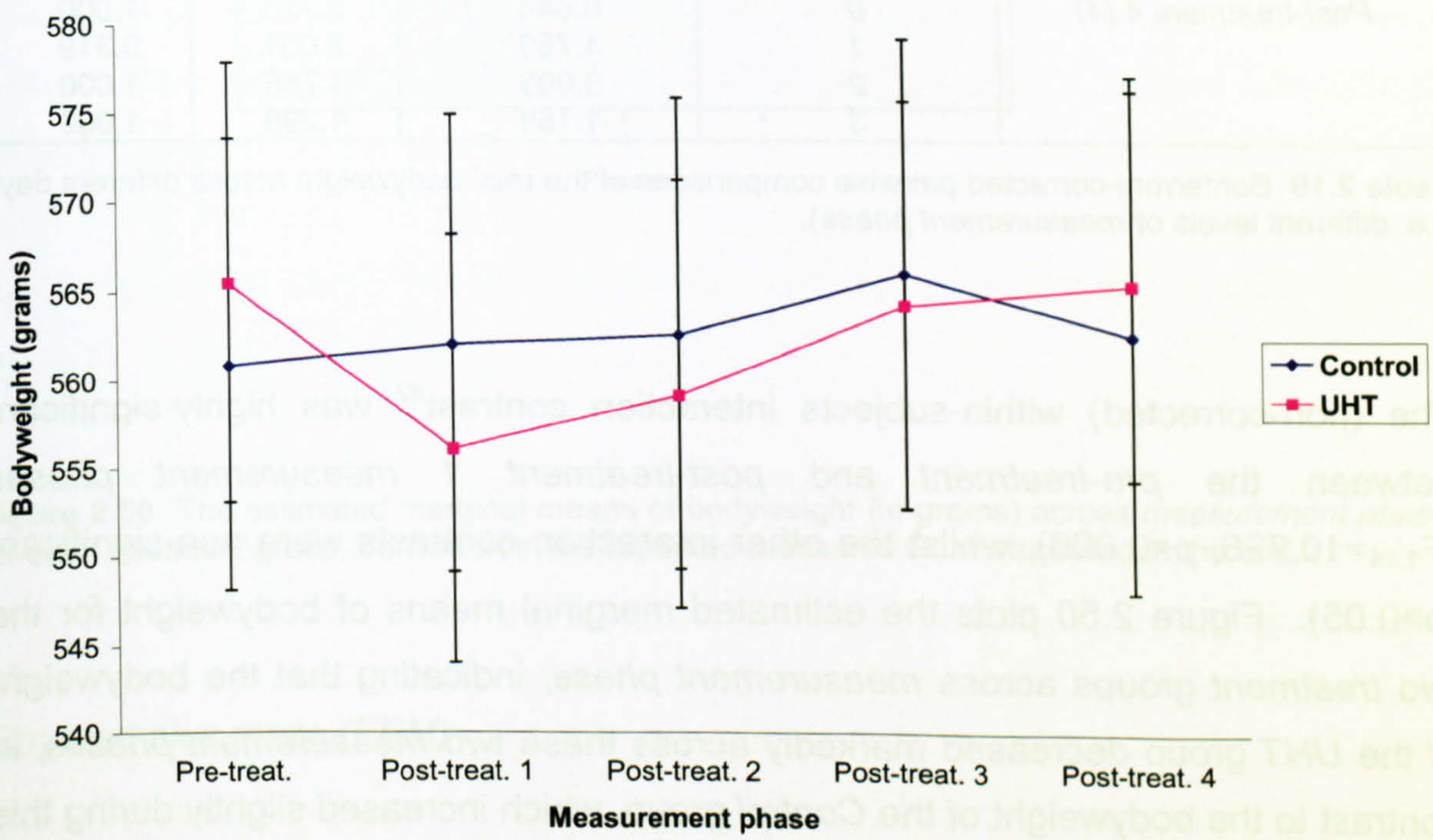

Figure 2.49 The mean bodyweight, in grams, by *treatment* group across *measurement phase* (weighed on one occasion prior to the *treatment*, and four occasions after the *treatment* had started; +/- 1SEM).

[81] Note: the data summarised here includes a rat in the *Control* group who suffered ill-health for a period between the final two days of bodyweight measurement (as mentioned earlier, this rat was excluded from the progressive ratio task as a result): this rat had the greatest percentage decrease in bodyweight across this period, although the *Control* group's mean bodyweight in the final *measurement phase* is still slightly lower than the *UHT* group's when this rat's data is excluded.

Measurement phase (I)	Compared with… (J)	Mean Difference (I-J)	Std. Error	*p*
Pre-treatment (0)	*1*	4.125	1.611	0.226
	2	2.363	2.466	1.000
	3	-1.813	2.189	1.000
	4	-0.644	2.725	1.000
Post-treatment 1 (1)	*0*	-4.125	1.611	0.226
	2	-1.763	1.380	1.000
	3	-5.938	1.280	0.004 **
	4	-4.769	2.001	0.319
Post-treatment 2 (2)	*0*	-2.363	2.466	1.000
	1	1.763	1.380	1.000
	3	-4.175	0.823	0.002 **
	4	-3.006	1.766	1.000
Post-treatment 3 (3)	*0*	1.813	2.189	1.000
	1	5.938	1.280	0.004 **
	2	4.175	0.823	0.002 **
	4	1.169	1.298	1.000
Post-treatment 4 (4)	*0*	0.644	2.725	1.000
	1	4.769	2.001	0.319
	2	3.006	1.766	1.000
	3	-1.169	1.298	1.000

Table 2.19 Bonferroni-corrected pairwise comparisons of the rats' bodyweight across different days (i.e. different levels of *measurement phase*).

The (non-corrected) within-subjects interaction contrast[82] was highly-significant between the *pre-treatment* and *post-treatment 1 measurement phases* ($F_{1,14}$=10.726, p=0.006), whilst the other interaction contrasts were non-significant ($p>0.05$). Figure 2.50 plots the estimated marginal means of bodyweight for the two *treatment* groups across *measurement phase,* indicating that the bodyweight of the *UHT* group decreased markedly across these two *measurement phases,* in contrast to the bodyweight of the *Control* group, which increased slightly during this period.

[82] The 'repeated contrasts' option was chosen, in which each *phase* is compared against the one preceding it.

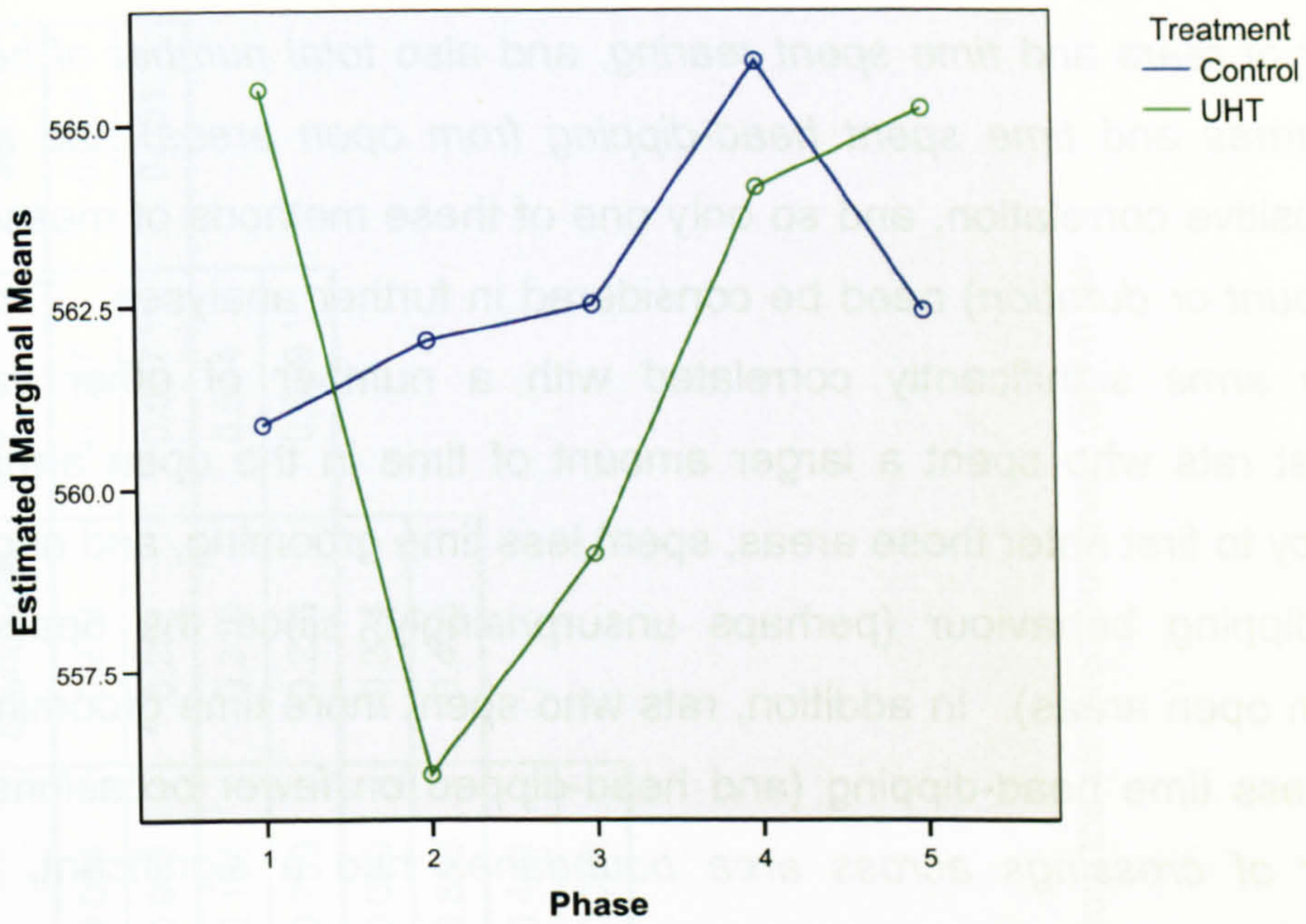

Figure 2.50 The estimated marginal means of bodyweight (in grams) across *measurement phase* for each *treatment* group, derived from the repeated-measures GLM discussed in the text.

Elevated plus maze (EPM)

Prior to analysis by *treatment*, a correlation matrix of several of the main variables of interest was first examined[83], namely: *time spent in open arms*; *latency to first enter an open arm (seconds); (log-transformed) time spent grooming; total number of crossings across area boundaries; time spent rearing; total number of head dips*

[83] We inspected this matrix to ensure we did not subsequently analyse a number of different variables which were correlated, and which therefore may be measuring the same, or similar, underlying behavioural construct (e.g. exploration, 'anxiety', etc.); plus, in a test of finite duration (5 minutes), if performing behaviour A means not performing behaviour B, some are likely to be correlated anyway.

from open areas; and *time spent head-dipping from open areas*[84]. As Table 2.20 shows, there were a number of significant correlations. Perhaps predictably, the *count* and *duration* data pertaining to the same behaviours (i.e. *total number of rears* and *time spent rearing,* and also *total number of head dips from open areas* and *time spent head-dipping from open areas)* had a highly-significant positive correlation, and so only one of these methods of measurement (i.e. either *count* or *duration*) need be considered in further analyses. *Time spent in the open arms* significantly correlated with a number of other variables, indicating that rats who spent a larger amount of time in the open arms had a shorter latency to first enter those areas, spent less time grooming, and engaged in more head-dipping behaviour (perhaps unsurprisingly, since the head-dipping occurred from open areas). In addition, rats who spent more time grooming spent significantly less time head-dipping (and head-dipped on fewer occasions). The *total number of crossings across area boundaries* had a significant, positive correlation with the *head-dipping* measures. Finally, the variables measuring *rearing* were not significantly correlated with any of the other variables in the matrix.

Following this inspection of the correlation matrix, the following variables were submitted to independent-samples *t*-tests, with *treatment* as a between-subjects factor*: time spent in open arms,* the *total number of crossings across area boundaries*, and the *time spent rearing.* The *time spent in open arms* neared significance at the 0.05 level (t_{14}=2.087, p=0.056), with the UHT group spending more time in this area of the EPM (see Figure 2.51); however, there was no significant difference in the *total number of crossings across area boundaries* (t_{14}=1.216, p=0.244), nor in the *time spent rearing* (t_{14}=1.570, p=0.139).

[84] The variables which measured duration (e.g. *time spent in open arms*) were expressed as a percentage of total session time (which was the same for each *subject*). The *total number of grooms* was not included, as the data could not be normalised.

		Time spent head-dipping from open	*Total no. of head dips from open*	*Time spent rearing*	*Total no. of rears*	*Total no. crossings across boundaries*	*Time spent grooming (log-trans.)*	*Latency to first enter an open arm*
Time spent in open arms	*r*	0.735	0.871	-0.038	-0.034	0.278	-0.647	-0.754
	p	0.001 **	<0.001 **	0.887	0.902	0.297	0.007 **	0.001 **
Latency to first enter an open arm	*r*	-0.535	-0.745	-0.052	0.101	-0.319	0.544	
	p	0.033 *	0.001 **	0.848	0.709	0.228	0.029 *	
Time spent grooming (log-trans.)	*r*	-0.715	-0.702	-0.060	-0.050	-0.224		
	p	0.002 **	0.002 **	0.825	0.853	0.405		
Total no. crossings across boundaries	*r*	0.593	0.630	0.397	0.441			
	p	0.015 *	0.009 **	0.128	0.088			
Total no. of rears	*r*	0.055	0.144	0.890				
	p	0.839	0.594	<0.001 **				
Time spent rearing	*r*	0.021	0.114					
	p	0.937	0.675					
Total no. of head dips from open	*r*	0.872						
	p	<0.001 **						

Table 2.20 Correlation matrix featuring a number of variables from the elevated plus maze (EPM) test.

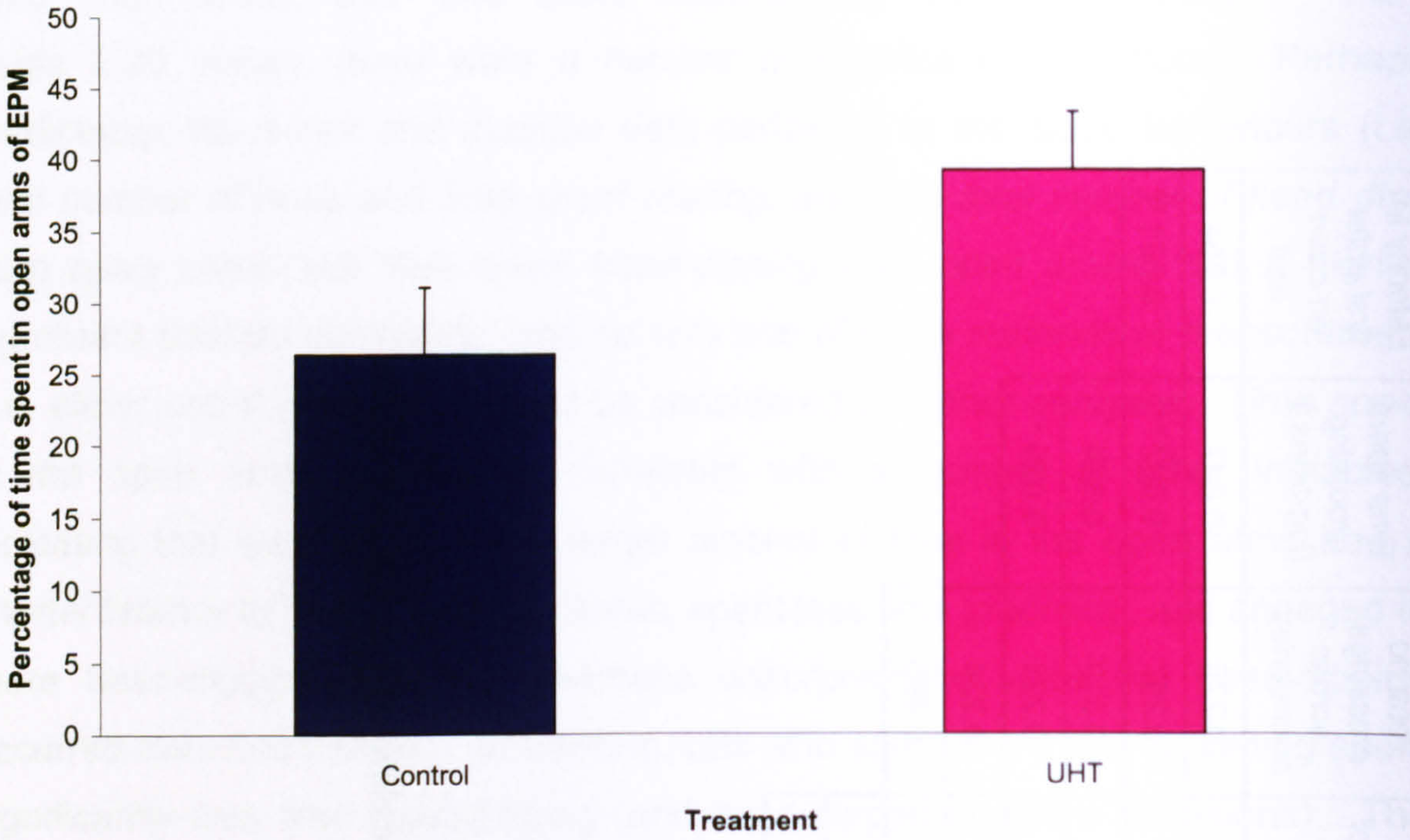

Figure 2.51 The mean percentage of test session time spent in the open arms of the elevated plus maze (EPM), by *treatment* (error bars = 1SEM).

Open field

Prior to analysis by *treatment*, a correlation matrix of several of the main variables of interest was again examined, namely the *(log-transformed) time spent in the central area of the arena; total number of areas entered; (square-root-transformed) time spent grooming; total number of rears; time spent rearing*[85].

As Table 2.21 shows, there were a number of significant correlations. Again, as one might expect, the *count* and *duration* data pertaining to *rearing* had a highly-significant positive correlation. In addition, the *time spent in the central area of the arena,* the *total number of areas entered,* and the *time spent rearing* were all highly-significantly correlated with each other, indicating that rats who spent more time in the *central area* of the open field arena also spent longer *rearing* (although the relationship with the *total number of rears* was not significant), and entered a

---

[85] Again, the variables which measured duration (e.g. *time spent in open arms*) were expressed as a percentage of total session time (which was the same for each *subject*). The *total number of grooms* was not included, as the data, again, could not be normalised.

significantly greater *number of areas;* the *total number of rears* had a highly-significant correlation with the *total number of areas entered*. Of these variables, the *total number of areas entered* was chosen for further analysis by *treatment*, since the *time spent in the central area* had a poor spread of values (mean: 1.46% of test session time, with values ranging from 0 – 5.73%; i.e. a likely floor/ceiling effect), and otherwise the correlations suggest it may be tapping similar behavioural tendencies (perhaps of exploratory activity) as the measurements of *rearing* – in which case only one variable need be analysed further. The *time spent grooming* was not significantly-correlated with any of the other variables, and so this was also submitted to further analysis.

Independent-samples *t*-tests indicated that the *total number of areas entered* was significantly-greater for the UHT group (t_{14}=2.586, p=0.022; see Figure 2.52), whilst there was no significant difference in the *time spent grooming* (t_{14}=-1.218, p=0.243).

		Time spent rearing	*Total no. rears*	*Time spent grooming (square-root-trans.)*	*Total no. of areas entered*
Time spent in central area (log-trans.)	*r*	0.714	0.460	0.127	0.628
	p	0.002 **	0.073	0.640	0.009 **
Total no. of areas entered	*r*	0.703	0.762	-0.038	
	p	0.002 **	0.001 **	0.888	
Time spent grooming (square-root-trans.)	*r*	0.215	-0.124		
	p	0.424	0.647		
Total no. rears	*r*	0.671			
	p	0.004 **			

Table 2.21 Correlation matrix featuring a number of variables from the open field test.

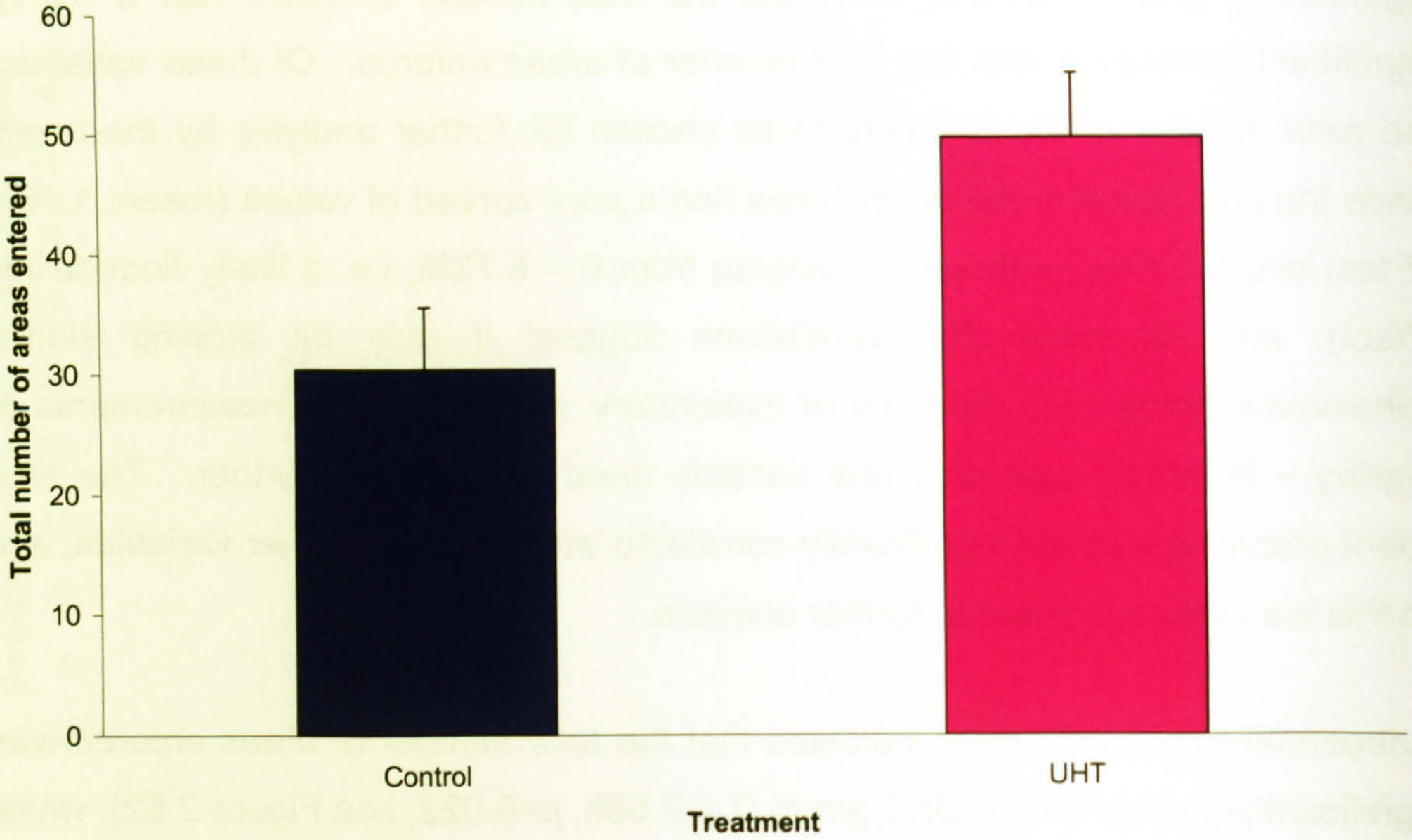

Figure 2.52 The mean number of floor areas entered in the open field test session, by *treatment* (error bars = 1SEM).

Addition of novel object to open field arena

An independent-samples *t*-test found that the *Control* group spent a significantly larger percentage of test session time in the peripheral area of the test arena (and thus less time in the central area) following the introduction of the novel object into the central area of the arena (following reflection, then square-root transformation: t_{14}=2.854, p=0.013; see Figure 2.53).

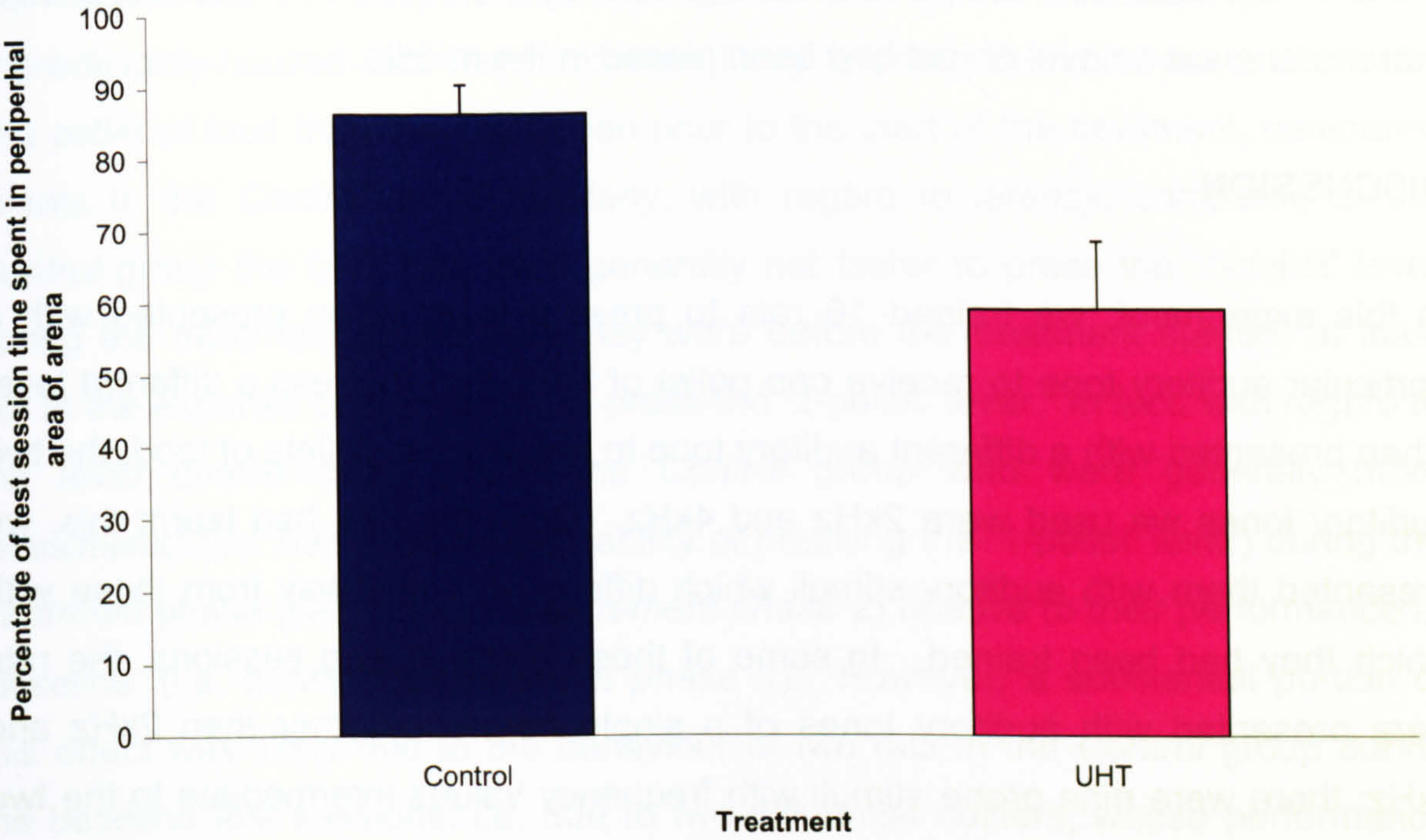

Figure 2.53 The mean percentage of test session time spent in the peripheral area of the test arena following addition of a novel object into the central area, for each *treatment* group (error bars = 1SEM).

Summary of concurrent test analyses

There were no significant effects of *treatment*, either as a main effect, nor as part of interactions, in the *Time taken to eat 50 pellets of food*, the *Lever-based progressive ratio test with food reinforcement*, nor in either of the *Sucrose preference tests*, although, in some, there were main effects of *measurement phase*. Curiously, the results from the two tests of 'food motivation' (i.e. the *Time taken to eat 50 pellets of food*, and the *Lever-based progressive ratio test with food reinforcement*) were significantly correlated, but in an *opposite* direction to that predicted by good concurrent validity.

The *bodyweight* data indicated that the *UHT* group lost a significant amount of weight at the start of the *treatment*, compared to the *Control* group; their weight then steadily increased as the *treatment* progressed, to ultimately recover its pre-treatment value. Finally, there was a near-significant trend for the *UHT* rats to spend more of their test session time in the open arms of the elevated plus maze, and they were significantly more active, in terms of their movement around the

arena, in the open field test, and spent significantly more time in the central area of that arena once a novel object had been placed in the middle.

DISCUSSION

In this experiment, we trained 16 rats to press a lever when presented with a particular auditory tone to receive one pellet of food, and to press a different lever when presented with a different auditory tone to receive two pellets of food; the two auditory tones we used were 2kHz and 4kHz. Once the rats had learnt this, we presented them with auditory stimuli which differed in some way from those with which they had been trained. In some of these probe-testing sessions, the rats were presented with auditory tones of a single frequency other than 2kHz and 4kHz: there were nine probe stimuli with frequency values intermediate to the two training tones, two with values less than 2kHz, and a further two with frequency values greater than 4kHz. In the other probe sessions, we presented probe stimuli consisting of both training tones played together. Half the rats then underwent a treatment intended to induce a negative change in their affective state: namely a series of unpredictable housing events designed to be mildly stressful; for the remaining rats, in the control group, the husbandry regime remained as it had done before (i.e. relatively 'predictable'). After 19 days of this treatment, all rats underwent testing with the probe stimuli once more. In all the probe-testing sessions, we recorded the response (lever) choice the rats made when presented with the different auditory stimuli, and also recorded their latency to do so. Since, in an pilot earlier study, we had established that rats prefer to receive two pellets of food over one pellet of food, we hypothesised that subjects undergoing the unpredictable housing treatment, designed to induce a negative change in affective state, would be more likely to judge ambiguous stimuli as having a relatively negative significance, or outcome: i.e. would be more likely to respond to the probe stimuli *as if* judging them to be the tone associated with the least-preferred outcome (by pressing the lever associated with one pellet of food). In addition, we hypothesised that these rats would be slower to press the lever which had putatively more positive associations (i.e. that associated with two pellets of food), and faster to press the lever which had putatively more negative (or less positive) associations (i.e. that associated with one pellet of food), relative to controls.

Our results did not support our hypotheses: during the treatment phase, the unpredictably-housed rats were *not* more likely to press the lever associated with one pellet of food than they had been prior to the start of the treatment, compared to rats in the *Control* group; similarly, with regard to *latency*, compared to the *Control* group the *UHT* rats were generally not faster to press the '1-pellet' lever during the *treatment* phase than they were before the treatment started, at least not at the expense of the latency to press the '2-pellet' lever. In fact, with regard to the *lever choice* data, it was the *Control* group who were generally more 'pessimistic' (i.e. had a higher probability of pressing the '1-pellet' lever) during the *treatment* phase (i.e. during *measurement phase 2*) relative to their performance at 'baseline' (i.e. during *measurement phase 1*). However, a substantial portion of this effect was likely due to the behaviour of two rats in the *Control* group during the baseline test sessions: i.e. due to two influential outliers, whose performance before any treatment had begun was markedly different from that of the other 14 subjects. *A posteriori,* it's difficult to partial out any outlying influence these rats had on the statistical models, at least not without making a number of subjective judgements which might generally compromise the veracity of the analyses, but it's certainly possible they may have masked other differences of interest in the data, and we can attempt to make some informed judgements to that end; we will address this later, along with a more detailed discussion of the rats' latency to record a response.

Concurrent tests

First, though, we will turn to the concurrent tests, since these revealed a number of noteworthy differences across *treatment*, which are worth discussing on their own terms, but also may help frame our subsequent discussion. Whilst the tests of sucrose preference and 'food motivation' (i.e. the *time taken to eat 50 pellets,* and the *progressive-ratio test*) revealed no significant differences across *treatment*[86], the results concerning the elevated plus maze, and the open field arena (before

[86] Although a test of concurrent validity between these two tasks revealed a curious negative correlation, which we shan't pursue further here, but which casts some doubt on the validity of at least one of these tests, as a measure of food motivation.

and after the introduction of a novel object), and also regarding changes in bodyweight across the *treatment* phase, were much more notable.

The rats in the *UHT* group crossed a significantly greater number of area boundaries in the open field arena[87], spent significantly more time in the central area of that arena once a novel object had been placed in the middle, and had a near-significant trend to spend more of their test session time in the open arms of the elevated plus maze. Both the elevated plus maze, and the open field test used in this study involve the enforced confrontation of the experimental subject with a novel, heterogeneous environment, certain sections of which offer more cover than others. The novel object test differs somewhat from this format, in that the test arena is no longer novel (relatively-speaking: i.e. the subjects have been in the open field arena for ten minutes prior to the introduction of the novel object), and so an unfamiliar (novel) element is introduced into an environment which is increasing in familiarity. As such, these tests are often used as indicators of 'anxiety-like' states and traits, with the assumption that an 'anxious' animal will have a greater tendency to avoid exposure (i.e. seek cover), and also a greater tendency to stay away from novel elements: i.e. will spend less time in the open arms of the EPM, less time in the centre of the open field arena, and less time in close proximity to the novel object (e.g. Belzung & Le Pape, 1994; Carobrez & Bertoglio, 2005; Ohl, 2003; Prut & Belzung, 2003; van Gaalen & Steckler, 2000). Thus, to the extent that these assumptions hold, the behavioural profile of the *UHT* rats in this experiment suggest an *anxiolytic* effect of the treatment, indicating that the *UHT* rats are less 'anxious', at least at the time of testing, that the *Control* subjects; these findings would, in general, run contrary to that predicted from a treatment designed to induce a negative change in affective state.

So, to what extent do these assumptions hold? Firstly, it's worth noting that, *a priori,* predictions regarding the effect of 'anxiety' on subjects' behaviour in these tests are not always obvious (e.g. Paul et al., 2005): for example, we might predict an 'anxious' animal would more actively seek escape from the test apparatus,

[87] This variable, in turn, had a highly significant positive correlation with the percentage of test session time spent in the central area of the arena.

perhaps spending more time on the open arms (from which escape is possible, if potentially hazardous) (e.g. Holmes et al., 2000), or would more actively engage in information-gathering and risk assessment by visiting and assessing exposed areas (e.g. Garcia et al., 2005). More generally, whilst the extent of psychopharmaceutical validation (i.e. with anxiolytics and anxiogenics), particularly with respect to the EPM and the open field test, is considerable, there are notable inconsistencies (e.g. Carobrez & Bertoglio, 2005; Hogg, 1996; Prut & Belzung, 2003), and it is possible that some of these empirical non-sequiturs reflect the diversity of *a priori* predictions, and more generally the complexity of 'anxiety' itself: behaviourally, cognitively, physiologically and genetically (e.g. Rodgers, 1997). Indeed, tests such as the EPM, open field and novel object, have been employed as means to diverse ends: for example, as measures of depression, locomotory activity, arousal, emotionality, emotional reactivity, neophobia, exploration, and so on (e.g. Belzung & Le Pape, 1994; Gronli et al., 2005; Harris et al., 1997; Kalueff & Tuohimaa, 2004; Kelley, 1993; Maslova et al., 2002; Ohl et al., 2001; Roth & Katz, 1979; Roy & Chapillon, 2004; Strekalova et al., 2005). Of course, there will be plenty of investigators who would debate how appropriate some of these tests are as measures of the variables I have just listed, but this diversity of application nevertheless strongly suggests that these tests may be sensitive to constructs which differentially map onto 'anxiety', some of which may be orthogonal to it.

Measures such as the EPM, open field and novel object test have been used in experiments employing treatments similar[88] to that used in the current study, and a survey of some of the results from these experiments will be instructive with regard to noting any precedents, or otherwise, for our findings, and is more generally useful in illustrating some of the issues, of prediction and interpretation, that we have been discussing.

For example, whilst some such studies have reported that 'unpredictable stress' treatments have had an effect on subjects' behaviour typically interpreted as

---

[88] By similar, we mean treatments consisting of a series of different events whose mode of delivery has an unpredictable element (in terms of time of delivery and duration), each of which is designed to be stressful.

'anxiogenic' (e.g. Maslova et al., 2002; Zurita et al., 2000)[89], others have found no effect (e.g. Gouirand & Matuszewich, 2005; Matuszewich et al., 2007; Mineur et al., 2006; Mitra et al., 2005; Vyas & Chattarji, 2004)[90], whilst others still have reported the opposite: i.e. an 'anxiolytic-like' effect of these treatments on subjects' behaviour (e.g. D'Aquila et al., 1994; Gronli et al., 2005; Harris et al., 1997)[91]. Maslova et al (2002) attempted to account for this variability with reference to stressor severity, treatment duration, the interval between the end of the treatment and time of testing (they suggested there was a more 'anxiogenic' behavioural profile as each of these variables increases), the age of the rodents, and so on. As this implies, there is typically a wide range of such differences between studies, and it is reasonable to cite them when attempting to reconcile contrary results, but acknowledging such variation in experimental design is different from providing causal and functional explanations which are predictive of behavioural change, and given the variability in design and results between studies, this is a very difficult objective to achieve. For example, the studies which Maslova et al cite suggest unpredictable stress treatments exert an increasingly 'anxiogenic' effect the longer they are applied (i.e. the longer the duration of the treatment phase), yet Strekalova et al (2005), running mice exposed to treatment phases of differing duration in an EPM test, found the *opposite* effect: i.e. the longer the treatment phase, the more 'anxiolytic' the effect.

---

[89] Zurita et al (2000) found that adult rats exposed to a seven day period of variable stress spent significantly less of their test session time in the open arms of an EPM, compared to a control group, when tested 7 days following the end of the treatment. Similarly, Maslova et al (2002) found that at the end of a 12-day period of unpredictable, stressful events, rats spent significantly less of their test session time on the open arms of an EPM, compared to a control group (this anxiogenic profile persisted when tested on the EPM a number of months later); incidentally, if rats were instead repeatedly *handled* over that 12-day period, they spent significantly *more* time on the open arms.

[90] Matuszewich et al (2007), Mineur et al (2006), Mitra et al (2005), and Vyas & Chattarji (2004), all used the EPM, whilst Gouirand & Matuszewich (2005) used an open field test. Incidentally, whilst Gouirand & Matuszewich (2005) found no significant difference between treatment groups in the time spent in the centre of the open field arena, nor in number of areas in the arena entered, they did find that rats undergoing an 'unpredictable stress' treatment reared more frequently, as did Harding (Harding, 2002; Harding et al., 2004).

[91] D'Aquila et al (1994) found that rats exposed to a chronic mild stress procedure for several weeks spent significantly more time in the open arms of an EPM than a control group, whilst having significantly lower sucrose intake (a putative test of 'anhedonia'); furthermore, whilst administration of a psychopharmaceutical with an anxiogenic action (picrotoxin) reduced the amount of time control rats spent in the open arms, it had no such effect on the experimental group. Gronli et al (2005) found rats that subjected to a chronic mild stress treatment for 4½ weeks crossed significantly more squares in an open field arena, compared to rats in a control group, with a non-significant tendency to cross more of the central squares; they suggested that this difference reflected the "psychomotor agitation" which can be associated with human depression. Harris et al (1997) found that rats receiving a chronic mild stress treatment for approximately seven weeks had a shorter latency to leave the starting corner, crossed more squares (both central and peripheral) and reared more frequently than a control group in an open field arena.

Similarly, Harris et al (1997) predicted that rats exposed to a chronic mild stress treatment would develop a 'depressive-like' state with a concurrent lowering of exploratory activity in an open field test (perhaps due to a decrease in exploratory 'interest', or similar); however, they found the opposite: rats exposed to a 'chronic mild stress' treatment had a shorter latency to leave the starting square, entered more inner and outer squares, made more rears, and so on. They speculated that the stressors they employed may have been too mild to induce a depressive-state, but noted that the significant weight loss in their experimental group suggested the treatment was indeed stressful, and so suggested that the "stress desensitized the animals in a novel environment when they were placed in the open field apparatus". However, it's not immediately obvious why apparent stress in one experiment increases exploratory behaviour, but has no effect in another (e.g. Gouirand & Matuszewich, 2005), and more generally, it's not especially clear why a treatment designed to induce a 'depressive-like' state would not induce co-morbid 'anxiety' (e.g. Zurita et al., 2000), with a more thigmotaxic behavioural profile (e.g. Heisler et al., 1998; Simon et al., 1994).

As this discussion suggests, the treatment effects we observed in the EPM, open field, and novel object test are notable and interesting, but their biological significance is not obvious. Nevertheless, we can attempt to narrow down a few *a posteriori* hypotheses, and weigh their relative plausibility.

There is evidence that putatively stressful extraneous (i.e. treatment-related) events can increase locomotory activity in subsequent novel situations, such as an open field arena (e.g. Roth & Katz, 1979). Indeed, when the putative stress of an 'anxiety' test, such as the EPM, is manipulated, for example by increasing its level of illumination[92], there is evidence that this induces a 'hyperlocomotion' confound in chronically-stressed mice, not found in a control group (Strekalova et al., 2005). Therefore, might the *UHT* rats in the current experiment simply be more active, at least within these test situations? With regard to their behaviour in the open field

[92] The EPM test used in the current experiment employed illumination levels which were (to our eyes) quite similar to that employed in the rats' homecage environment, using red lamps of the same wattage. In contrast, Streklova et al (2005) found 'hyperlocomotory' effects only when using white illumination of a minimum of 25 lux: considerably greater than that employed in the current experiment.

arena, there are certainly grounds for judging that this may be so. However, when looking at the suite of tests as a whole, the proportional differences between the two treatment groups, in the time spent in various sectors of the test apparatus, suggests that such an explanation is perhaps unlikely to hold on its own. For example, when the novel object is introduced into the open field arena, the *UHT* rats spend proportionately more time in close proximity to it, and they also have a tendency to spend proportionately more time in the open arms of the EPM[93]; i.e. whilst they may be more generally active, their profile of activity differs qualitatively, not just quantitatively.

Some studies, including some of those briefly outlined above, suggest that animals who have been exposed to stressful events are more 'robust' in the face of further stress (this phenomenon, or a variant of it, is sometimes referred to as 'stress inoculation'; e.g. Fox et al., 2006). For example, there is evidence that extraneous (i.e. treatment-related) stress can have a 'protective' (classically 'anxiolytic-like') effect when confronted with the putative, acute stress of an 'anxiety' test, such as the EPM. Haller & Halasz (1999), for instance, found that rats who had been previously group-housed, but were isolated for five days prior to testing, exhibited an 'anxiogenic' behavioural profile in an EPM, spending less time in the open arms when compared to rats who had continued to be group-housed. However, when rats undergoing this 'isolation stress' treatment were subjected to additional daily social defeats, this 'anxiogenic' effect of isolation was abolished. Similarly, Morato & Brandão (1997) found that when previously group-housed rats were isolated, or when group-housed rats were exposed to a series of novel situations prior to testing, they spent less time in the open arms of an EPM, compared to controls. However, when these putative stressors were combined (i.e. isolated rats exposed to novelty) the effect was 'anxiolytic', with the rats spending more time on the open arms. It's possible, therefore, that experiencing putatively intense stress 'toughens one up', to put it colloquially (and rather unsatisfactorily), when faced with further

[93] NB as outlined in the Results section, there was no significant difference in the total number of arm entries (i.e. *total number of crossings across area boundaries*); furthermore, although not reported in that section, there was also no significant difference between the two *treatment* groups in the number of closed arm entries (t_{14}=0.617, p=0.547). These two variables are often used as indicators of general locomotory activity in the EPM (e.g. Cruz et al., 2005; File, 2001; Hogg, 1996; Morato & Brandao, 1997).

stress, but such an observation has little predictive power: how much stress is enough? Should it vary quantitatively, qualitatively, or both?

If we refine this hypothesis a little further, it may become more plausible. In the current study, the *Control* group were kept largely undisturbed in their homecages, much as they had been, bar operant training and testing, since their arrival in the laboratory. The *UHT* rats, on the other hand, had undergone a number of weeks of a treatment involving both lone occupancy of a variety of novel apparatus, and repeated handling by an experimenter: characteristics shared with the EPM, open field, and novel object test situations. If the *UHT* rats had become somewhat habituated to such manipulations, i.e. learning that their consequences are *relatively* mild and finite, then the difference in stress induced by these tests (i.e. the EPM, open field, and novel object), compared to their day-to-day experiences, may be substantially less for the *UHT* group than it is for the *Control* rats, for whom these manipulations are largely without precedent. Whilst the 'unpredictable' nature of such treatments is designed, in part, to ensure against any habituation or adaptation (e.g. Cabib, 1997), there are, nevertheless, predictable elements to them: in the current treatment regime, perhaps milder than many employed elsewhere (e.g. Banasr et al., 2007; Harkin et al., 2002; Willner et al., 1987; Yalcin et al., 2007), each type of event recurs, they all occur during the dark phase, they all occur in test cages of the same design, and so on.

With regard to the data pertaining to *bodyweight*, the analysis indicated that the *UHT* group lost a significant amount of weight at the start of the *treatment*, compared to the *Control* group, but then their weight steadily increased as the *treatment* progressed, to ultimately recover its pre-treatment value, and indeed to exceed the bodyweight of the *Control* group[94]. Again, the relationship of putative stress with changes in bodyweight is, empirically, somewhat equivocal (e.g. Forbes

[94] NB There is a suggestion, from the summarised data, that the *UHT* rats subsequently gain weight, following initial loss, at a rate greater than that of the *Control* group. Whilst this was not found to be significant, it's perhaps worth noting that a significantly greater rate of weight gain in an experimental group, compared to a control group, has been found elsewhere (e.g. Harkin et al., 2002, employing an unpredictable stress treatment, which didn't include food or water deprivation, with mice).

et al., 1996; Harkin et al., 2002; Steinberg & Watson, 1960)[95], but stress is *typically* associated with a lower rate of weight gain, or indeed weight loss (e.g. Broom & Johnson, 1993; Harris et al., 1998; Marti et al., 1994; Santos et al., 2000).

The pattern of bodyweight change observed in the current experiment - of initial loss, and then recovery, in the face of putative stress - has been observed in a variety of other studies, and it may be useful to look at these in a little more detail. Steinberg & Watson (1960), for example, introduced a different husbandry event (each putatively stressful) every eight days to rats in a 'disturbed' treatment group, leaving the control group 'undisturbed'; following each event, the 'disturbed' rats underwent an acute loss of bodyweight, but this loss recovered towards the end of each eight-day period, despite the persistence of the putative stressor. This recursive pattern continued, as each new event was introduced, until no more changes were made to their husbandry regime (i.e. the rats were 'undisturbed'), at which point their rate of weight gain recovered to match that of the control group. Similarly, O'Connor & Eikelboom (2000) found that moving rats from isolated to paired-housing induced an acute period of initial weight loss, although the rate of weight gain soon recovered to match previous values; likewise Marin et al (2007) reported that rats submitted to a variable stress treatment (which, incidentally, included periods of food/water deprivation) underwent an initial weight loss, but their rate of weight gain recovered to match controls, even when the treatment persisted.

If we assume, for a moment, that a reduction in the rate of weight gain is a (very) rough proxy of perceived stress (e.g. Broom & Johnson, 1993), the changes in bodyweight in the current experiment, and also in the studies just discussed, suggest that these treatments may be perceived as particularly stressful at their onset, but the level of perceived stress attenuates as the animal adjusts to the new regime. To take a physiological example, Pfister (1979) found that

[95] For example, Steinberg & Watson (1960) found a reduction in net weight gain, relative to controls, in rats subjected to a successive sequence of putative stressors, whilst Forbes et al (1996) found no effect of a chronic stress treatment on weight gain in rats, and finally Harkin et al (2002) found that mice subjected to repeated unpredictable stress gained weight at a greater rate than controls; none of these treatments involved food or water deprivation.

glucocorticosterone[96] levels were initially elevated in rats placed in a novel environment for 30 minutes, but as this treatment was repeated, daily, hormonal levels reduced to a level comparable to controls; i.e. a degree of physiological habituation seemed to have taken place.

So, in summary, a number of the concurrent tests suggested that the unpredictable housing treatment had an 'anxiolytic-like' effect on the rats, and/or was generally 'arousing', or 'activating'. To the extent that a change in bodyweight is a very rough proxy of perceived stress, the data suggests that the treatment may have been initially stressful, but the level of perceived stress abated as the rats adapted, or habituated, to it; this, perhaps in turn, may have been reflected in the 'anxiolytic' profile of the rats in the *UHT* group.

Affect-related predictions: re-visiting the human literature

As mentioned earlier, the unpredictable housing treatment employed in this study was designed to induce - via the administration of unpredictable stressors - a negative change in affective state, perhaps akin to a 'depression-like' mood, and/or an 'anxiety-like' state. As our introductory chapter suggests, the relationship of affective states and traits with cognition and behaviour, at least in humans, is multifaceted and complex, but for our experimental hypotheses, we selected specific predictions concerning the co-variance of affective state with responses to ambiguity and uncertainty, and designed the experiment to this end. As briefly summarised earlier, the treatment we employed to try and induce a negative change in affective state did not change the subjects' responses in the direction we hypothesised. Of course, it's possible that our treatment induced no such change in subjects' affective state, or it may have been that the specific behavioural indices we chose to focus on when gauging the level of empirical support for our hypotheses were insensitive to real biological change. To help us address these two possibilities, and before returning our discussion to the operant tasks which formed the central part of the experiment, it's worth first widening the scope of our

[96] A hormone associated with stress responses (e.g. Moberg & Mench, 2000).

predictions regarding the possible effect of stress and negative affective change on cognitive-behavioural functioning.

Firstly, if we assume that the unpredictable housing treatment did induce a 'depressive-like' state (e.g. Willner et al., 1987), and further assume that certain characteristics of such a state are shared with depression in humans (e.g. Willner, 1997), then we may well predict dysfunction in a range of cognitive processes, some of which are likely to be relevant to performance in an operant discriminatory task such as the one employed in the current experiment. In fact, depression in humans has been found to be associated with impairments in "every cognitive domain" (Gualtieri et al., 2006), including attention (e.g. Keilp et al., 2008; Mialet et al., 1996), memory (e.g. Burt et al., 1995; Ebmeier et al., 2006), a number of aspects of executive control and functioning (e.g. Biringer et al., 2005; Stoddart et al., 2007), and, more generally, the slowing of a variety of psychomotor processes (i.e. a decrease in the speed of mental and motor functioning; e.g. Pier et al., 2004; Schrijvers et al., 2008). So, as we discussed in our opening chapter, in humans at least, depression is not only associated with the characteristic *biasing* of certain cognitive processes, but also a decline in their *capacity*, or *efficacy*. As one might predict, the extent of such impairment depends on the severity and type of depressive disorder, but even those with mild depression, and no complicating factors, tend to have impaired cognitive functioning compared to controls (Gualtieri et al., 2006). Such cognitive impairments are typically attenuated following successful anti-depressive treatment (e.g. Gualtieri et al., 2006), and remission (i.e. recovery) from a depressive episode (e.g. Biringer et al., 2005).

If the unpredictable housing treatment we employed did induce a 'depressive-like' affective state in the rats subjected to it, and this was in turn associated with a degree of cognitive impairment, as it is in humans, what effects might we predict this would have on responding in the operant task we employed? Well, we may well predict a slower reaction time: Gualtieri et al (2006), for example, found depressives had a longer response latency than controls in certain neurocognitive tasks, although only those involving a considerable amount of cognitive effort. Den Hartog et al (2003) found an effect of depression on response speed, but this time only in neurocognitive tasks involving relatively *automatic* cognitive processing,

rather than more complex, effortful paradigms. As this suggests, there is some empirical disagreement as to the circumstances necessary to elicit differences in response latency, and of course these disagreements have implications theoretically (i.e. with regard to the mechanisms underlying such differences in behavioural output), but the general observation holds (in these, and other studies: e.g. Kalb et al., 2006; Murphy et al., 2001; Rose & Ebmeier, 2006): depression in humans is generally associated with slower responses in cognitive-behavioural tasks, and we might therefore predict that an analogous affective state in rodents would be associated with similar changes in response latency in a task such as the one employed in this experiment. We might also expect to see a reduced level of accuracy: certainly, depression has been associated with an increase in error rate in a variety of cognitive tasks (e.g. Rubinsztein et al., 2006; VollmerConna et al., 1997), although it's important to note that other studies have found no effect of depression on accuracy (e.g. Murphy et al., 2001), or even an *improvement* (at the expense of response speed: e.g. Rose & Ebmeier, 2006). Of course, the presence, or otherwise, of differences in error rates across affective state will likely depend on the nature of the task employed to measure them, as well as the type and severity of the depression, and so on; but in general, though, given the range of cognitive impairments which can be found in depressed people – e.g. in problem-solving abilities, cognitive 'flexibility' (including initiation of behaviours and perseveration), and so on (e.g. Gualtieri et al., 2006) – it is reasonable to expect that accuracy of performance will be compromised in certain circumstances.

Soon, we will apply these *a posteriori* predictions to our empirical findings from the current experiment, but it's worth highlighting one particular observation in this literature, which is of more general concern to experiments employing paradigms such as the one used here: namely, that relating to a deficit in various *inhibitory* processes. Kaiser et al (2003), for example, employed a task in which subjects were asked to press a button each time lower-pitched tones were presented (which was often), but to refrain from pressing the button on the relatively rare occasions high-pitched tones were heard. They found that depressed patients had a deficit in 'response inhibition': i.e. they were more likely to incorrectly press the button compared to controls; however, when these contingencies were reversed (i.e. when they were only asked to press the button each time they heard the *relatively-*

rare high-pitched tones), they did not differ from controls. A greater tendency to make 'dominant', 'default', or 'pre-potent' responses is also associated with anxiety, and, more generally, with stress (Eysenck et al., 2007; Mendl, 1999); there is the possibility, then, that 'depressed', 'anxious' or 'stressed' animals may be more likely to perform such pre-potent responses, and if those responses are those operationalised as 'optimistic' (e.g. they are associated with a more valued outcome, and therefore, at baseline, are more likely to be performed), then we have a direct confound with respect to our experimental hypotheses.

Otherwise, with regard to cognitive-behavioural performance in 'anxious' people, Eysenck et al (2007) note that anxious subjects often perform with lower levels of *efficiency* in a variety of such tasks, although their *effectiveness* often remains the same. By this, they mean that response latency, for example, is often slower in anxious people, compared to controls, although the end result of their performance, in terms of accuracy, is often comparable. They relate this, in turn, to the manner in which cognitive resources are deployed in those who are anxious, suggesting that certain attentional faculties are cast more widely, maximising the field of sensitivity to detecting possible threat. This, in turn, results in fewer attentional resources focused on the ongoing task in hand, as indeed we discussed in Chaper 1 (although, of course, this rather depends on the task: the deployment of attentional resources would likely differ if the test itself had features which were perceived as threatening). As a result of a change in the functioning of mechanisms such as these, the *efficiency* with which neurocognitive tasks are performed is often compromised in those who are anxious, although *effectiveness* (e.g. accuracy) can be maintained by employing compensatory strategies, and greater effort. As mentioned above, another consequence of such differences in attentional control is a reduction in the ability to inhibit automatic, dominant or pre-potent responses, since executive attentional control processes, which are otherwise involved in such inhibition, are more widely deployed (Eysenck et al., 2007).

As we touched on earlier, in a wide-ranging review of the effects of stress on various aspects of cognitive functioning, Mendl (1999) noted a similarly increased tendency to make 'default' responses under conditions of stress. Otherwise, he

notes that the effect of putative stress on memory, learning, and decision-making processes is generally complex. For example, learning and memory functioning can be enhanced, or impaired, depending, in part, on the timing and nature of the stressor(s) with regard to the learning or mnemonic event. In general, though, when stress is chronic, the effect on learning and memory tends to be detrimental. Similarly, response latency in certain tasks can be increased (i.e. slowed) perhaps due to stress-related lapses in attention, although under some circumstances it can actually be *quicker*, although such an increase in speed is typically at the expense of response accuracy.

Alongside documenting such complexity in the relationship of stress with cognitive functioning, Mendl (1999) notes that there is a general prediction regarding the pattern of this relationship, which, whilst over-simplistic, is nevertheless a relatively enduring generalisation, typically referred to as the Yerkes-Dodson law. When plotted, this law describes an inverted U-shape, with some aspect of cognitive performance on the y-axis, and some indicator of stress on the x-axis; as such it predicts that 'under-aroused' animals will perform relatively poorer than 'moderately-stressed' animals, who in turn perform relatively better than 'greatly-stressed' animals: i.e. there is a level of stress, or arousal, which is optimal with regard to task performance, and this level is *not* at either terminal end of an (abstract) index of stress.

Post-hoc evidence for a widened range of affect-related predictions

Now that we've widened the scope of our predictions regarding the relationship of negative affective states and, more generally, stress, with aspects of cognitive-behavioural functioning, we can revisit our data to better gauge the actual effect of the treatment we used.

A number of these predictions concerned response latency: so, what effect did the treatment we employed have on this variable? Whilst the *UHT* group tended to be *overall* faster to record a lever-press response (i.e. regardless of which *measurement phase* they were in), this difference between the two *treatment* groups was significantly greater in the second *measurement phase*, both in the

single-frequency probe sessions, and in the training sessions which preceded them (i.e. the significant interaction of *treatment* with *measurement phase* was persistent across different operant task designs). This pattern was generally found when modelling *all* lever presses and trial types, but was especially apparent when modelling solely those presses made on the '2-pellet' lever (in the single-frequency probe sessions), and the '2-pellet' reference trials (in the training sessions). As outlined above, a relatively-enduring prediction from the literature relating to human subjects is for response latency to be lengthened as affective state becomes more negative (specifically, more depressed and/or anxious); in the current experiment, not only do we find no such trend, we actually find just the opposite. However, as Mendl's (1999) review indicated, a quicker response latency in putatively stressful circumstances is not without precedent, but this typically occurs at the expense of accuracy: i.e. there is a speed-accuracy trade-off. In certain circumstances, then, it appears that stress can induce animals to behave more *impulsively*[97]. So, is this the case with regard to the *UHT* rats? Well, despite the increase in the rapidity of their responding in the training sessions, there is no decline in the *accuracy* of their responding (in fact, there was a slight, non-significant, improvement in accuracy, relative to controls; e.g. see Figure 2.13). Elsewhere, whilst the *UHT* group were not significantly faster to record a response in the dual-frequency probe sessions, there was a significant interaction between *treatment* and *measurement phase* when modelling the accuracy of responding to the reference trials which occurred in those sessions, with the *UHT* group more accurate, and the *Control* group less accurate, in the second *measurement phase* compared to the first. Finally, although responding to the reference stimuli (i.e. 2kHz and 4kHz) was not analysed in isolation in the single-frequency probe sessions, there was little indication that the *UHT* group were less accurate as they became faster to record a response: e.g. if we look at Figure 0.15, and Figure 0.16, in the Appendix, which plotted predicted responding from a model in MLwiN which allowed *lever choice* to vary across higher-order interactions which included *probe value, treatment, measurement phase* and *contingency*, we see that whilst there is some suggestion

[97] Evenden (1999) notes that "the concept of impulsivity covers a wide range of 'actions that are poorly conceived, prematurely expressed, unduly risky, or inappropriate to the situation and that often result in undesirable outcomes'"; as such, it can be operationalised as the biasing of the speed-accuracy trade off in favour of the former (e.g. Berlin et al., 2004).

that the *UHT* rats in the *2kHz=2pell contingency* group were a little less accurate with regard to responding to the '2-pellet' reference tone in the second *measurement phase*, compared to the first, this is not particularly the case with regard to the '1-pellet' reference tone, and moreover, the *UHT* rats in the *4kHz=2pell* group are substantially more accurate in *phase 2*, compared to *phase 1*, when presented with the reference stimuli.

Similarly, earlier, we made passing comment (in a footnote) on a paper which had suggested that the increase in exploratory activity they had observed in rodents exposed to a putatively stressful treatment may have been akin to the 'psychomotor agitation' found in some human depressives (Gronli et al., 2005). Such an "activated", "excited" state of psychomotor agitation is listed as one of the possible criteria leading to a diagnosis of depressive disorders (e.g. *Diagnostic and Statistical Manual of Mental Disorders IV*, 1994), is found in a substantial portion of depressed patients[98], and may contribute towards a discrete subtype of depression (e.g. Akiskal et al., 2005). To the extent that the *UHT* rats in the current experiment seem relatively more 'activated', and changes in locomotory activity and related indices as a result of chronic stress has been documented elsewhere, including in a number of the studies already discussed, is it possible that the treatment we employed induced a similar 'agitated' state? Again, it is the increase in efficiency, and effectiveness, in the responding of the *UHT* rats which suggests this is not the case: psychomotor agitation, at least in humans, is characterised, among other things, by "racing thoughts" and "risky behaviour" (e.g. Olgiati et al., 2006) - traits unlikely to lead to focused and accurate behavioural output in a relatively-demanding cognitive-behavioural task.

So, we are left in a position in which our experimental hypotheses have not been supported, and a number of the other predictions we generated after further considering the possible effects of stress, and/or a negative change in affective state, on responding in tasks such as those we have employed have also not been realised. Instead, the overall picture is one which is counter-intuitive, with those

[98] For example, Olgiati et al (2006) state that psychomotor agitation it is reported in 20-30% of outpatients with major depression, and in 70% of inpatients.

rats submitted to the unpredictable-housing treatment more *efficient* (in terms of their latency), and more *effective* (in terms of their accuracy). In addition, whilst we encountered some outlying individual variation at 'baseline' which was likely to be particularly influential with regard to the findings of the model, our findings suggested that, if anything, it was the *UHT* group who were more 'optimistic' regarding the post-treatment biasing of their responses.[99] In addition, as we have already discussed, much of the decrease in response latency (i.e. quickening of response speed) that occurred in the *UHT* group, across *measurement phase,* was specific to the '2-pellet' lever, and the '2-pellet' reference trials. Furthermore, in the single-frequency probe sessions, the increase in response speed with regard to the '2-pellet' lever for the *UHT* rats was especially apparent for those *probe values* far from the '2-pellet' reference tone: again, a pattern in keeping with that predicted from an 'optimistic bias'. Finally, the maintenance of accuracy with regard to the reference trials, despite this 'optimistic' tendency, argues against the possible confound of 'stressed' animals, or those in a relatively negative 'affective state', being more likely to perform 'pre-potent', 'dominant', 'default' responses.

Given the substantial impact of two *subjects* in the *Control / 2kHz=2pell* group on the statistical models we employed, there is, of course, the possibility that our findings have simply been skewed, either by the particularly unusual responding of those two rats, or by other individual variation of a substantially outlying nature we have otherwise failed to detect. However, the concurrent tests we employed strongly suggested that our *treatment* had a detectable physiological, and behavioural, effect, and it is therefore reasonable to conclude that the changes we observed across *measurement phase,* in the operant tasks we conducted, were genuinely *treatment*-related.

---

[99] For example, repeated-measures ANOVAS conducted on various aspects of probit functions fitted to the single-frequency probe data, reported in Appendix C, found that the point of probable bisection (i.e. the intermediate probe value where there was an esimated even chance of pressing either lever) was generally in an 'optimistic' direction across *measurement phase* for all *treatment / contingency* sub-groups bar the *Control / 2kHz=2pell* group. However, when analysing the fitted probability of responding to the intermediate *probe value* closest to the '2-pellet' reference tone, the significance of the interaction between *treatment* and *measurement phase* indicated that differences was more pervasive across *treatment,* with the *UHT* group generally more 'optimistic' in the second *measurement phase* than they were in the first, compared to the *Control* group, who were generally more 'pessimistic'.

In fact, the only prediction which concurs with our findings is that pertaining to the Yerkes-Dodson law, with the *UHT* rats somewhere towards the peak of the inverted-U the law describes, and the *Control* rats further down the slope of the curve, perhaps towards the end of the x-axis signifying lower levels of stress and arousal; i.e. the putative stress the *UHT* rats have undergone has *enhanced* their performance in the operant task, in terms of *efficiency,* and *effectiveness.* As Mendl (1999) notes, the Yerkes-Dodson law is an over-simplified picture of actual response to stress, and, as a purely-descriptive account, is agnostic with regard to any underlying mechanism, but nevertheless it is a surprisingly pervasive generalisation.

Affect-related change in operant performance: conclusions

So, we're left with the somewhat suprising possibility that the *treatment* we employed *enhanced* the functioning of the rats subjected to it, *may* have rendered them more 'optimistic', and (operationally-speaking) emboldened them in the face of the putative anxiety induced by a variety of the concurrent measures we employed. Is that a reasonable position to defend? As an enriching, welfare-enhancing husbandry intervention, the unpredictable-housing treatment would have very poor face validity, but there may be elements of it which, when contrasted against the regime the *Control* rats were subjected to, may, on some indicators, have had a positive impact. Meehan & Mench (2007) discuss the 'enhancing' effects an intermediate level of stress can have on functioning, noting that "not just high stress due to intense or prolonged aversive events, but low stress due to inadequate basic stimulation, can result in maladaptive behavioral and physiological responses". This somewhat implies that such a relationship may occur across *quantitative* variations in stress, and this very well may be so, but a considerable portion of their review actually concerns *qualitative* variation: specifically, those stressors which are challenging, yet can be coped with or even mastered, those in which the contingencies are relatively explicit, and those over which animals can exert a degree of control, may, on some indices of welfare, induce positive change. It is possible there were subtle elements of the *treatment* we employed which met some of these criteria. As mentioned earlier, whilst an important element of treatments such as these - i.e. those designed to induce

negative changes in affective state through the administration of putatively stressful husbandry events - is their unpredictability, the *treatment* we employed had elements which were nevertheless predictable. Certainly, the data pertaining to *bodyweight* change suggests a degree of physiological (and perhaps behavioural, with regard to feeding) adaptation took place as the *treatment* progressed. In addition, the probe-testing in the current study began later in the *treatment* phase (on Days 19-22) than the probe-testing conducted by Harding et al (2004) in their experiment, employing a similar treatment (probe-testing took place between Days 10-19 of the treatment in their study); therefore, if habituation and adaptation does occur as such treatments persist, the extent of this may have been greater in the present experiment. With regard to the intervention which was, *a priori,* perhaps likely to be the most challenging - namely the introduction into a unfamiliar conspecific's homecage whilst that conspecific was present - the social stimuli rats were older than the experimental rats, and thus closer to the end of the expected lifespan of a laboratory rodent. As such, the experimental rats may have had the better of them in any aggressive encounters, gaining a degree of mastery and control over an otherwise stressful situation (e.g. Haller & Halasz, 2000).[100] Incidentally, this contrasts with the treatment employed by Harding et al (2004), in which the social stimuli animals were of very similar ages.

Elements of the *a posteriori* narrative we are beginning to construct are clearly speculative, and indeed we could continue noting differences between the design of the current experiment and that of comparative studies which have found results at odds with our own, seeking meaning with the benefit of post-hoc hindsight; more generally, though, we are on firmer ground in relating our results to the relatively well-established phenomenon that a 'non-zero' level of stress, somewhere intermediate to 'higher' and 'lower' values on an abstract index of 'arousal' is associated with optimal cognitive performance.

[100] Note, as discussed in the Method section, these events were terminated whenever *damaging* aggression was seen to occur. Whilst such incidences were very rare, some aggressive interactions were observed, although records were not kept of the specifics of such encounters (i.e. of submissive behaviour, etc.)

Intentions, habits and the extent of training

Later in this thesis, we will describe an experiment which is a more explicit test of what *subjects* actually 'know' about the consequences of their actions in the operant tasks we have designed; the answer to that question is, of course, important, since it has some heavy bearing on the veracity of our experimental hypotheses. However, by revisiting certain aspects of the rats' behaviour, both as they learnt the operant task, and then as they subsequently engaged with it from a position of relative expertise, we may be able to glean some useful insights into the matter now. To this end, it is some way reassuring that over the first few training sessions the rats received, their errors were biased very strongly in favour of the lever associated with receipt of the larger quantity of food (e.g. see Figure 2.11); as suggested by our earlier pilot study (see Appendix A, p.310), this clearly shows that the subjects were sensitive to the difference in reward size, and had, operationally, a preference for the larger reinforcer. Of course, since all subjects reached the criterion that we had set for them to be considered to have fully learnt the task – i.e. to consistently respond above chance on *both* reference trials – this bias in errors abated. However, although not reported in the Results section, it is striking that the tendency to commit such asymmetric inaccuracies persisted: half of the rats failed a binomial test with regard to the '1-pellet' trials in the next training session they received following attainment of criterion[101]; for many (six of these eight), this was on the following day.

In some senses this is a little disconcerting: in operant training procedures such as this, the implicit assumption is generally that once animals reach such a criterion, we have reassured ourselves that they will carry on performing at that pitch, and they can then proceed to the next stage of the experiment. However, whilst some kind of marker (be it a performance criterion, or a discrete number of trials) is a necessary component of such experimental designs, the location of this marker is,

[101] Note, all the rats, including these eight, performed significantly above chance with regard to the accuracy of responding to the '2-pellet' reference trials.

of course, arbitrary[102]: in this instance, for example, it did not mean that the subjects would thereafter perform as they had done over the three days in which they satisfied the criterion, although they were considerably more accurate than they had been a number of sessions prior to that attainment. As we have discussed, though, in another sense such asymmetric inaccuracies *are* reassuring, for the following reasons: in adopting a discriminatory operant procedure, this, and similar studies (e.g. Harding et al., 2004; Matheson et al., 2008), have had to make a number of choices regarding the specifics of that design: duration of inter-trial interval, number of training sessions, length of training sessions, etc, which may interact in complicated ways with the respective treatments adopted. Given the breadth, and the complexity, of the literature pertaining to operant learning theory, we can do no more at this juncture other than note that, of course, these parameters may be important, but one well-established phenomenon is worth sketching out in a little more detail, as it potentially blunts the sensitivity of these sorts of experiments to detecting real biological signals. As Dickinson, Balleine, and their colleagues, have noted in a number of papers (e.g. Balleine & Dickinson, 1998; Dickinson & Balleine, 1995; Dickinson et al., 1995), the *extent* of operant training can be an important determinant of the significance of the relationship, for an animal, between a stimulus (e.g. a lever), the response (pressing the lever), and its consequence (delivery of food). Specifically, they have found that more extensive operant training is associated with increasing *insensitivity* to *incentive learning*. For example, Dickinson et al (1995) found that when rats that had been trained to press a lever to receive food reward in training sessions (n=120 rewarded trials) following a period of food deprivation were subsequently exposed to the 'operant food' outside the operant chamber in conditions of satiation (i.e. when they had had free access to their standard 'lab chow'), they subsequently reduced the extent of their lever presses when later placed back in the chamber, but again only when satiated. If they *hadn't* had access to the 'operant food' in conditions of satiation outside of the operant chamber, they continued to press the lever with the same frequency as unsatiated controls even when they had free

[102] 'Arbitrary' in similar sense as defining a significant p-value as under 0.05 is arbitrary: obviously it's an educated judgement to place the marker there (rather than at, say, 0.40), but of course there's nothing *intrinsically* significant about the value, other than the fact we have historically decided to name it so.

access to lab chow (i.e. were satiated) prior to the operant session. That, in itself, is a very interesting finding, suggesting reward devaluation is conditional on having prior experience of being satiated when presented with the reinforcing food, but more pertinently for our purposes, rats who received three times the number of training trials (i.e. n=360 rewarded trials), prior to the treatments we have just described, were insensitive to such attempts at reward devaluation: they simply kept on pressing at the same (relatively-high) rate no matter what their previous experience of satiation and 'operant food' exposure. This suggests that, after extended training, actions which are intentional, and goal-directed, insofar as they are sensitive to changes in the value of that goal, become more habitual: i.e. they are better characterised by a learnt association between the stimulus (e.g. a lever) and the response (e.g. pressing the lever), without a nuanced appreciation of the consequences of that response.[103]

In the current experiment, the hypotheses centre around what an animal 'knows' or 'expects' the consequences of their actions to be: it is impossible to respond optimistically[104] when the consequences of all choices are undifferentiated. The discriminatory operant designs that both the current experiment, and similar studies, have adopted, obviously adds a layer of complexity absent from Dickinson & Balleine's operant design (in which the choice was to press a lever or not, not to choose between different levers following the presentation of different stimuli), so their experiment is not directly comparable to ours, but it is important to at least register concern that both our study, and that of others, have trained rats *beyond* an initial criterion, and this *may* have consequences as to the manner in which the animals conceptualise the goal of their actions[105]. In a similar vein, Phillips & Barr (1997) suggest "caution in the use of well trained operant responses to assess changes in reinforcement value or levels of incentive motivation in animal models of depression". It is in this important sense, then, that it is reassuring to observe

---

[103] Elsewhere, Weiskrantz (2001) has characterised this shift as from a behavioural pattern which is 'on-line', to one which is 'off-line'.

[104] In the folk psychology sense of the word, rather than "responding *as if* optimistic".

[105] As mentioned in the Methods section, problems with the building in which the rats were housed meant the probe-testing, and thus the start of the *treatment* phase, were delayed, forcing a longer training period than would otherwise have been adopted; so perhaps the concern is particularly strong here.

that just prior to probe-testing, in each of the *measurement phases*, the pattern of the rats' responding was consistent with a sensitivity to the consequences of their actions (e.g. see Figure 2.12).

The effects of contingency

In this Discussion, we have generally made very scant reference to the effect of *contingency*, other than noting that there were two *subjects* in one particular *contingency* group who had a particularly unusual pattern of responding in one of the *measurement phases.* This is despite the fact that *contingency* played a prominent role in a number of our analyses: of course, this was necessarily so, allowing us to appropriately partition any variance attributable to *contingency* so that the variance associated with the remaining factors in our model could be more accurately gauged. But in addition, it revealed some interesting effects of *contingency*, which we may, in turn, be able to relate to the likely psychophysical character of the stimuli we chose. For example, in the single-frequency probe sessions, the *2kHz=2pell contingency* group were overall less likely to press the '2-pellet' lever. Interestingly, this is what we would expect if, as noted in our Method, the perceptual differences between stimuli which differ by the same Hertz value were greater at lower Hertz frequencies: to the rats in the *2kHz=2pell* group, then, more of the probe stimuli appear more similar to the '1-pellet' reference tone (4kHz) than they do to the '2-pellet' reference tone (2kHz); hence, a response bias by *contingency*. As Figure 2.23 and Figure 2.24 suggest, the *4kHz=2pell* group are less *discriminating* across *probe value;* i.e. the slope of their response curve is shallower. This is presumably due to the asymmetry of reinforcement: if each reference stimulus were reinforced by the same quantity of food, the shape of the curves would (of course) be the same (bar individual variation). As a general point, if more of the probe stimuli *appear to be* the '2-pellet' reference tone for the *4kHz=2pell* group, then more of their '2-pellet' lever responses (they make more overall) are likely to be unrewarded; i.e. it is possible that the probability of reinforcement is lower, following responses to tones they perceive as heralding the larger quantity of food. Therefore, psychometric factors could have an important bearing on the effect of any *treatment* used in studies such as these (e.g. Matheson et al., 2008, who found substantial effects of contingency on patterns of

responding across probe value), and there is an argument for simply dispensing with this level of design complexity (i.e. by having only one contingency group, e.g. Bateson & Matheson, 2007), if one were confident that there was no asymmetry with regard to the psychophysical properties of the probe values employed across treatment (e.g. if one were confident that stimulus X was not perceived differently by animals in different affective states (perhaps due to changes in the rate of 'internal clocks', or the emotional significance of certain tones, etc.) *regardless* of any differences in reinforcement schedule).

Concluding remarks

If we assume, as we did initially, that the unpredictable housing treatment induced a negative change in the rats' affective state, then our results did not support our hypotheses. However, if we widen the scope of our predictions, and relate our results from the operant procedures to a number of the concurrent tests we conducted, then the picture becomes more complex, and we must entertain the possibility that our treatment induced a change in affect which was counter-intuitive. This is interesting in its own right, but more generally, whilst we made an informed choice when selecting our treatment, it is possible that alternative methods may enable us to realise our initial objectives, of affect-related change, in a manner which is less ambiguous. To this end, we conduct a similar study in the following chapter, but employ an alternative method to induce a change in affect.

Otherwise, it's worth noting that by considering a wider range of affect-related predictions, we gained some useful insights into the operant responding we recorded, and future studies may also benefit from incorporating such issues into a broader predictive and explanatory framework.

CHAPTER 3

'ENVIRONMENTAL ENRICHMENT' AND JUDGEMENTS OF AMBIGUITY IN RATS

INTRODUCTION

The current study follows closely on from the last chapter, again using operant discrimination tasks to measure rats' responses to apparent ambiguity across a treatment designed to induce a change in affective state. We employ a similar methodology - and indeed the same experimental subjects - as in the previous experiment, but on this occasion we attempt to manipulate the rats' affective state by varying the level of 'environmental enrichment' in their homecages; in addition, we also administer probe tests on more than one occasion during the treatment phase, to try and tap any changes in affective state as that phase progresses.

The provision of certain homecage materials (such as shelters, nesting substrates, ropes, etc.) is well-characterised as having a positive impact on indices of rodent welfare, stress, and putative affect, at least in contrast to more barren environments (e.g. Fox et al., 2006; Wurbel, 2001; Young, 2003). Indeed, the specific changes in 'enrichment' provision we employ in the current study are associated with positive (or negative, depending on the treatment group) changes in the welfare of rats as indexed by a variety of behavioural and physiological measures (Burman et al., 2006); in fact, this particular treatment, or very close to it, has been used in two recent experiments investigating changes in 'cognitive bias' across putative affective state, both of which found treatment effects in the hypothesised direction (Burman et al., 2008a; Burman et al., 2008b). Why might such changes in 'enrichment' provision be associated with changes in welfare, and affect-related indices? The provision of a shelter may provide space for a captive animal to retreat, perhaps allowing it to avoid or defuse antagonistic encounters, for example (e.g. Morgan & Tromborg, 2007). In addition, the provision of nesting material will likely allow rodents to satisfy at least some of their behavioural 'needs', or 'response rules' (e.g. Wurbel, 2001) – e.g. for nest construction – which may be otherwise thwarted if such substrates are not available. More generally, a

greater variety of substrates within the homecage may enable an animal to organise and structure its environment to a greater extent, perhaps increasing its sense of control over its captive surroundings (e.g. Fox et al., 2006).

As in the previous experiment, the current study employed a repeated-measures design, with operant probe-tests administered both before, and during, a treatment phase. At the start of the treatment phase, half the rats had some of their pre-existing 'enrichments' *removed* (*unenriched* group), whilst the remaining subjects received *additional* 'enrichments' (*enriched* group). In contrast to the previous study, instead of administering the probe-test sessions on just one occasion during the treatment phase, here we administered the probe-test sessions twice during the treatment: soon after the initial change in 'enrichment', and then a number of days later, after the rats had been exposed to the treatment for a longer period. As such, we get two 'snapshots' of behavioural responding as the treatment progresses. Indeed, the findings of the study described in the last chapter suggested that the unpredictable-housing treatment used in that experiment may have been particularly stressful when it was first introduced, but that this attenuated over time, despite the persistence of that treatment. Similarly, with regard to the current experiment, it's possible that the initial onset of the treatment for the *unenriched* group – i.e. the removal of modest 'enrichments' – may be the most stressful or distressing period, and that this enriched-to-unenriched contrast leads to an acute phase of negative affect (e.g. Burman et al., 2008b; Latham & Mason, 2006). Alternatively, it may be that the withdrawal of 'enrichments' induces a negative change in affective state only if exposure to it is longer-lasting (see the following for a discussion of the importance of 'enrichment' duration on a variety of neurological and behavioural parameters in rodents: Amaral et al., 2008; van Praag et al., 2000).

Since the rats used in the previous study were performing to a generally high level in the operant tasks in which they were trained, we decided to employ them again as subjects in this experiment, counterbalancing their previous treatment grouping across current treatment assignment, and additionally leaving a substantial rest period between the end of the previous study, and the start of the current one. As

such, we were able to further employ their trained skills, whilst reducing the overall number of experimental subjects used (*cf.* recruiting naïve rats).

Our experimental hypotheses remain the same as in the previous chapter: rats undergoing a treatment designed to induce a negative change in affective state - namely the withdrawal of environmental enrichments – are predicted to be less likely to respond to a variety of probe stimuli as if they presage the relatively better outcome of two pellets of food, compared to a treatment group given extra 'enrichments'. The *unenriched* treatment group are also predicted to be slower to respond as if anticipating the relatively better outcome (i.e. slower to press the lever associated with two pellets of food), and quicker to respond as if anticipating the relatively poorer outcome (i.e. quicker to press the lever associated with one pellet of food), than the *enriched* treatment group.

METHOD

Overview

Figure 3.1 provides a summary of the experimental schedule, highlighting the three *measurement phases* in which operant discrimination tasks were administered:

- *phase 1* – prior to the onset of the treatment phase;
- *phase 2* – early in the treatment phase;
- *phase 3* – late in the treatment phase.

The diagram further illustrates that a number of concurrent tests[106] were conducted at the end of the *treatment* phase. Hereafter we refer to the first day of the *treatment* phase as Day 1.

[106] NB since a number of the concurrent tests employed in the previous experiment, such as the EPM, open field, and novel object test, rely, at least partly, on their novelty to be effective tests of 'anxiety', exploratory activity, 'neophobia', and so on, we decided not to employ them again in the present study.

Treatment grouping (*enriched* or *unenriched)*, reference stimuli (2kHz or 4kHz), responses (left or right lever), quantity of food reinforcement (1 or 2 pellets), experimental room in which the operant training and testing took place, and *prior treatment* grouping (*UHT* or *Control*) were all counterbalanced in this experiment. As in the last chapter, hereafter *2kHz=2pell* refers to the *contingency* group in which correct responses to 2kHz tones were associated with receipt of 2 pellets of food, whilst correct responses to 4kHz tones were associated with reinforcement with 1 pellet of food; these contingencies were reversed for the *4kHz=2pell* group.

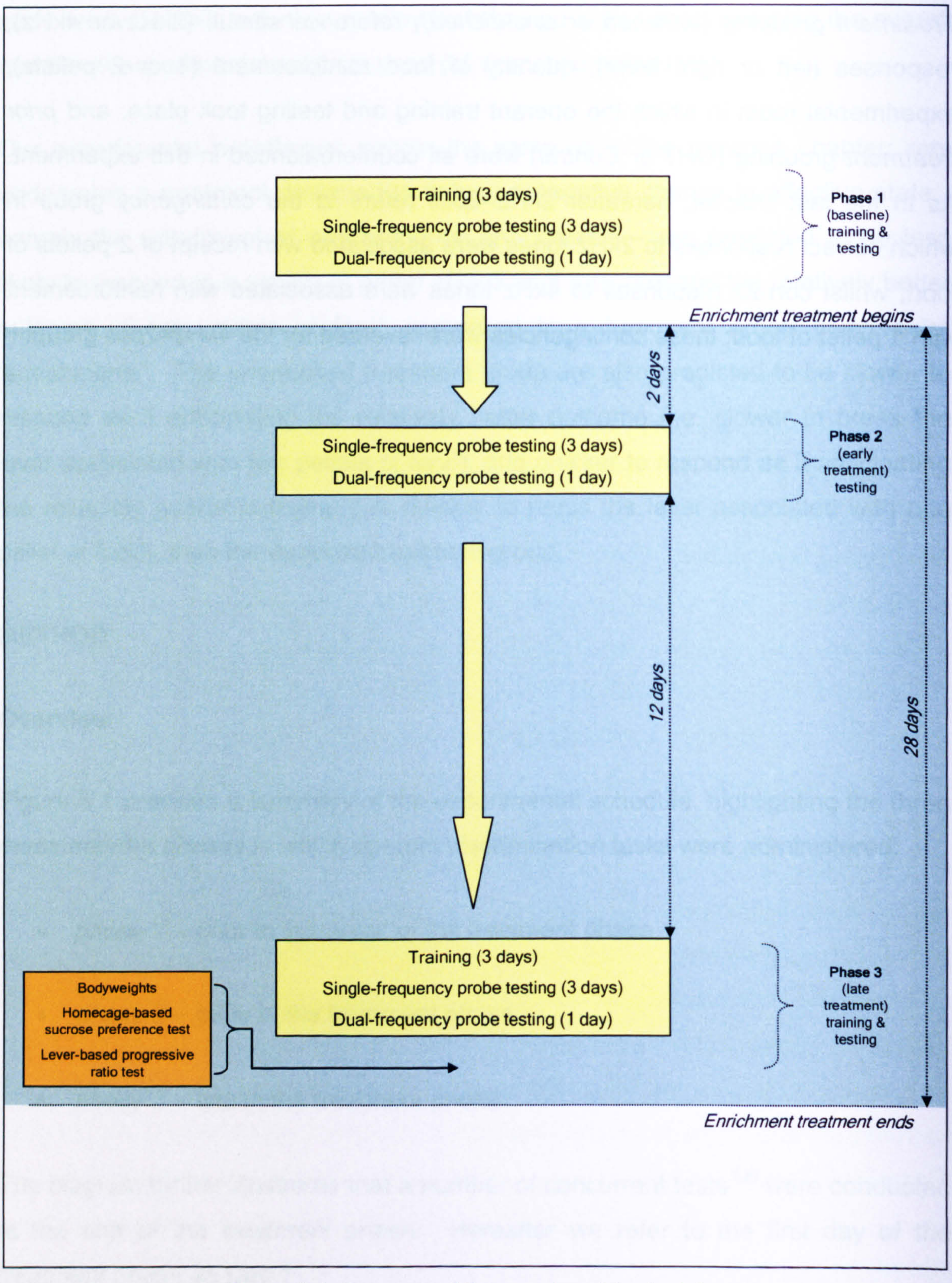

Figure 3.1 Summary of the experimental schedule.

Subjects and housing

The experimental subjects were 16 male Lister hooded rats (*Rattus norvegicus;* Harlan UK Ltd., Bicester, UK); they were previously used as *subjects* in the study described in Chapter 2, and by the time the treatment phase in the current experiment commenced, two months had passed since the termination of the treatment employed in the previous experiment.

The rats were housed in stable pairs, in cages measuring 56cm (L) x 34cm (W) x 19cm (H), with a 12:12 hour lights on:off cycle (lights off at 9am). Their homecages were cleaned on the same morning each week, contained sawdust bedding (Lignocel), and provided *ad libitum* access to food (Eurodent Diet 22%) and water. Before, and after, the *treatment* phase, their homecages also contained the following: shredded paper for nesting, a red Perspex shelter (Lilico, UK) and a chew block; the contents of the homecages differed from this *during* the *treatment* phase, and this is described on p.177.

The rats were checked daily for health throughout the experiment.

Two-choice operant discrimination training

Two-choice operant discrimination training, with differential reinforcement

The rats received one session of training per day, on the three consecutive days immediately prior to *phase 1* probe-testing, and on three days[107] immediately prior to *phase 3* probe-testing. The design of these training sessions was exactly the same as that described on p.59.

[107] Not consecutive: there was a gap of a day between the first, and second, day of training (for non-experimental reasons).

Single-frequency probe testing

The rats received one session of single-frequency probe-testing per day, three days in a row, in each of three different *measurement phases* (i.e. they received a total of nine sessions): the single-frequency probe-testing in *phase 1* (i.e. pre-treatment) finished the day before the *treatment* began, the probe-testing in *phase 2* took place on Day 3 of the *treatment*, whilst the probe-testing in *phase 3* took place on Day 23 of the *treatment*. The design of the single-frequency probe test sessions was exactly the same as that described on p.61.

Dual-frequency probe testing

The rats received one session of dual-frequency probe-testing, in each of three different *measurement phases* (i.e. they received a total of three sessions). Each dual-frequency probe-testing session took place the day after the final single-frequency test in that *measurement phase*. Therefore, the dual-frequency test in *phase 1* took place just before the start of the *treatment* phase (the change in 'enrichment' provision took place the same day, at the end of testing); the test in *phase 2* occurred on Day 6 of the *treatment*; and finally the test in *phase 3* occurred on Day 26 of the *treatment*. The design of the dual-frequency probe test sessions was exactly the same as that described on p.63.

Concurrent tests

Bodyweight

The rats were weighed, in counterbalanced order, on Day 27 of the *treatment*.[108]

[108] NB given the notable results from the analyses of bodyweight in the previous experiment, it would have been sensible to weigh the rats more frequently in the current study; I actually thought I *had*, and was surprised, on re-visiting my lab notes, to find that I had not done so more often (although, of course, their general health was checked daily throughout the study). It's possible that I may have misplaced some data pertaining to bodyweight measurement, but I think this is unlikely – otherwise, it was simply a (substantial) oversight on my part.

Homecage-based sucrose preference test

All rats underwent a sucrose preference test in their homecage, the design of which was exactly the same as that described on page 65, except that it terminated after 7 hours (i.e. earlier), due to the detection of possible ill-health in one of the rats (in the *enriched treatment* group). The test was conducted on Day 27 of the *treatment*.

Lever-based progressive ratio test with food reinforcement

The design of this test was exactly the same as that described on page 66. Since one of the *subjects* (in the *enriched treatment* group) was withdrawn from the experiment the day before (due to ill health, as mentioned above)[109], only 15 rats took part in this test, conducted on Day 28 of the *treatment* (the rats' homecage 'enrichment' provision was reverted to baseline levels at the end of this test).

Enrichment treatment

After the completion of probe-testing at 'baseline' (i.e. during *measurement phase 1*), the provision of 'enrichments' in the rats' homecages was altered, so that half the rats had fewer 'enrichments' than they had before (*unenriched* group), whilst the other half had more 'enrichments' than they had before (*enriched* group). This change in 'enrichment' provision lasted 28 days, after which the provision reverted to the pre-existing regime (see p.175).

The homecages of the *unenriched* group had the shredded paper, shelter and chew block removed, whilst the homecages of the *enriched* group had sisal rope[110] and two compacted cotton nestlets (Lilico, UK) added as extra 'enrichments' (e.g. Burman et al., 2008a; Burman et al., 2008b); see Figure 3.2.

---

[109] This *subject* was soon after euthanised, due to the detection of an inoperable tumour. In the attendant veterinary surgeon's opinion, the tumour was unlikely to have been causing the animal any pain or discomfort, and therefore is perhaps unlikely to have affected the rat's behaviour greatly in the operant tests in which he was a subject.

[110] One length, attached to the cage roof in three locations (at each end, and in the middle), forming two suspended loops.

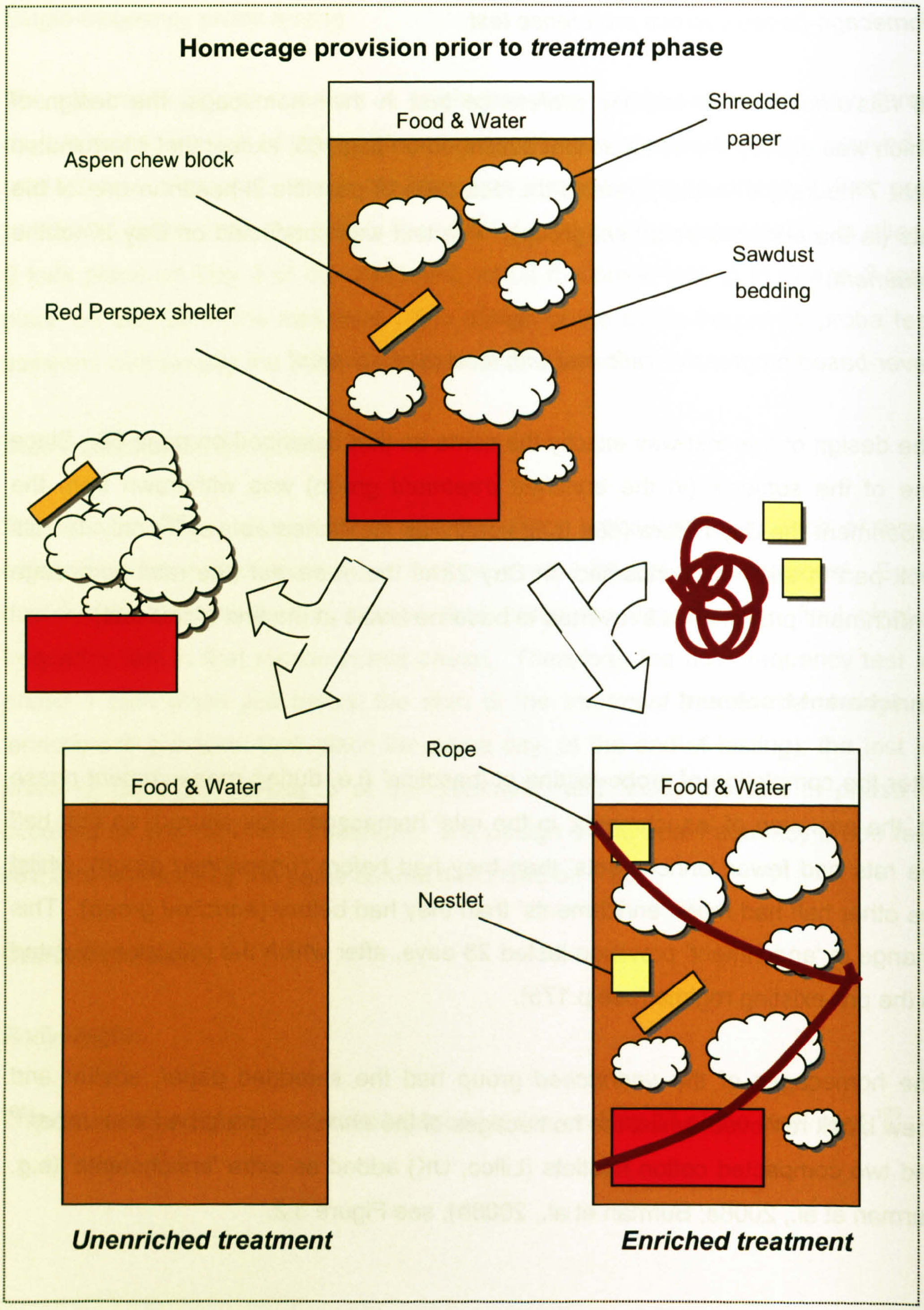

Figure 3.2 Schematic diagram illustrating the change in the homecage provision of 'enrichments' at the start of the *treatment* phase.

Data analysis

In the last chapter, we analysed the *lever choice* and *latency* data from the single-frequency probe sessions using both repeated-measures ANOVAs, in SPSS (presented in the appendices), and also in a multilevel analysis, using MLwiN. Whilst it is of some academic interest to compare and contrast different analytical approaches, one of the main reasons we conducted two types of analysis was to gauge whether the results of the more complex multilevel analyses, which may ultimately prove more useful to us, seemed reasonable given the results of the repeated-measures ANOVAs. That proved to be the case, and so now we focus only on multilevel analyses of the *lever choice* and *latency* data from the single-frequency probe sessions in this chapter. Furthermore, since we introduced our multilevel analyses in some detail in the last chapter, rather than reiterate that information, here we will only present the main findings. Otherwise, the significance tests, the estimation procedures, and the basics of model specification (i.e. selection of response (y) variables, centering of continuous predictor (x) variables), remain the same as in the previous experiment, as do our general methods of model-fitting. Once more, only non-reinforced trials, in which a lever press was recorded, are included in the analysis.

All other analyses were conducted in SPSS 14.0, and all met the assumptions of the statistical procedures used, except where we highlight an issue. When conducting repeated-measures ANOVAs with a within-subjects factor which had more than 2 levels (i.e. $k>2$), as in the last chapter, we follow the advice of Quinn & Keough (2002), and reject the null hypothesis if either the adjusted univariate output, or the multivariate output, reports significance at the 0.05 level. Unless otherwise stated, when reporting test statistics in such instances, we quote the Greenhouse-Geisser adjusted univariate output.

RESULTS

Training performance

For reference, Figure 3.3 plots the mean accuracy in the reference trials in those training sessions conducted just prior to probe-testing in the first (i.e. pre-treatment) and third *measurement phases* (i.e. towards the end of the *treatment*), indicating a general improvement in overall accuracy across the sessions in each of the two *measurement phases,* and a general bias towards making more mistakes in the reference trials associated with one pellet of food.

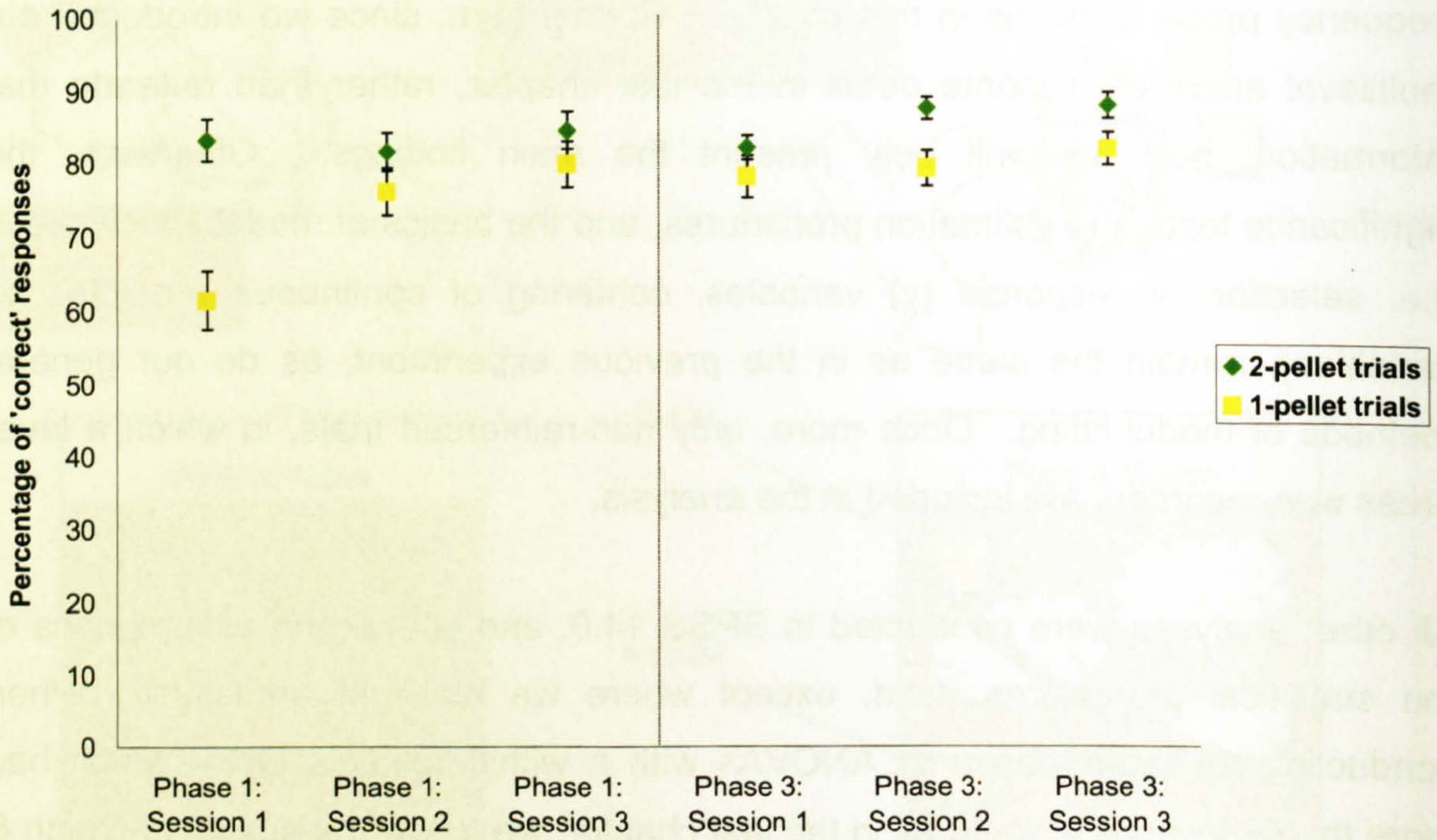

Figure 3.3 The mean percentage of 'correct' responses, across training session, for each of the two types of training trial, summarised by quantity of associated food reinforcement (+/- 1SEM). As indicated on the x-axis, the data is taken from each of the training sessions on the consecutive days immediately prior to single-frequency probe-testing in the first and third measurement phase.

Figure 3.4 plots the same data, but pooled across *measurement phase,* and summarised by *treatment* group. Analyses revealed that there were no significant differences between the *treatment* groups in the accuracy of responding to each type of reference tone, nor across all reference tones, either over the three training

sessions prior to probe-testing in *phase 1* ('2-pellet' trials (Y^4-transformed): t_{14}=1.193, p=0.253; '1-pellet' trials: t_{14}=0.890, p=0.389; all trials: t_{14}=0.182, p=0.858), nor over the three training sessions prior to probe-testing in *phase 3* ('2-pellet' trials (square-transformed): t_{14}=1.365, p=0.194; '1-pellet' trials: t_{14}=0.750, p=0.466; all trials: t_{14}=0.034, p=0.974).

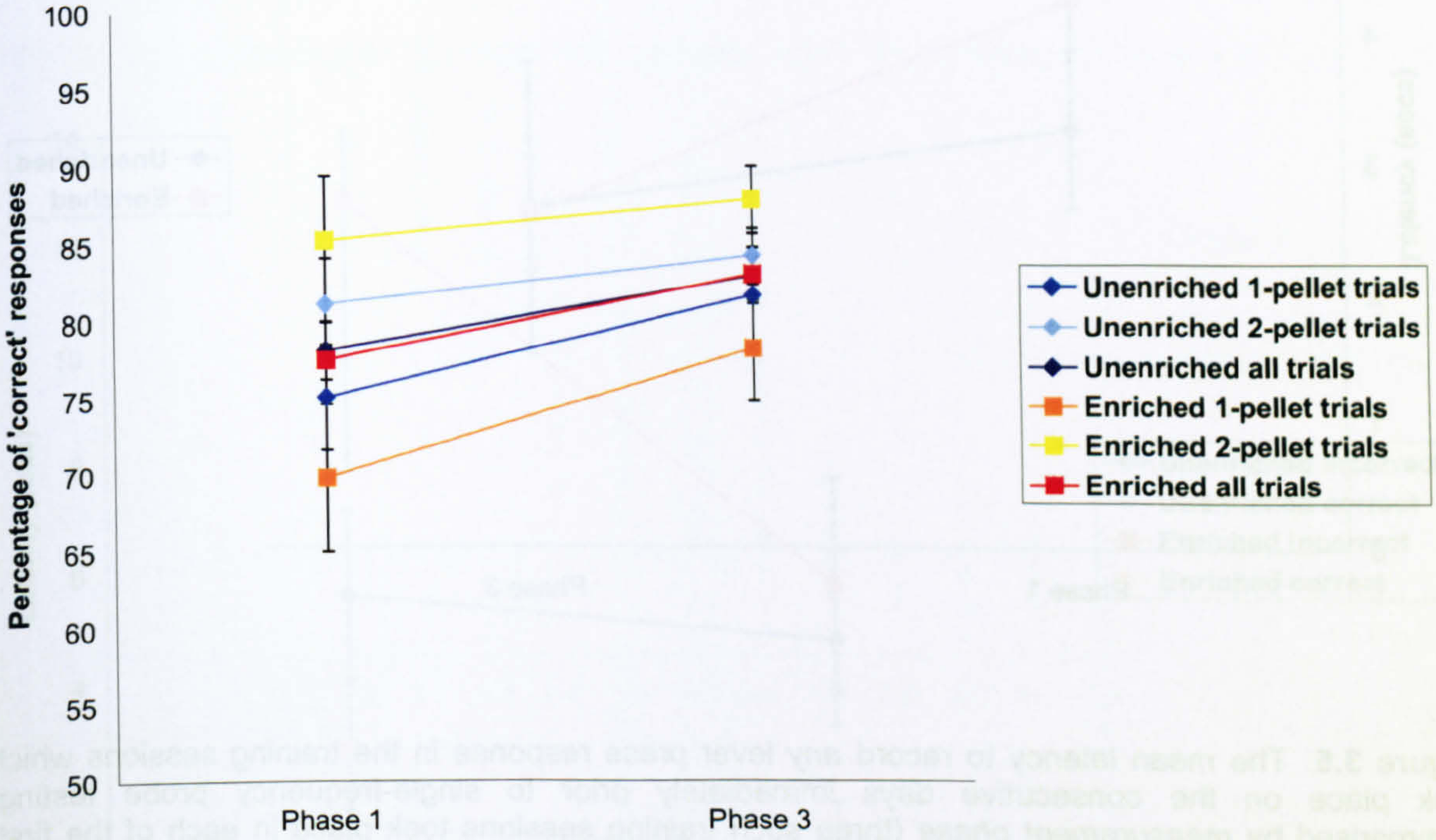

Figure 3.4 The mean percentage of 'correct' responses in the training sessions which took place in the first and third *measurement phases*, by *treatment.* The data is summarised by trial type (+/- 1SEM).

In addition, the mean latency to press a lever (of any identity) in the training sessions in the first and third *measurement phases* was submitted to a repeated-measures GLM, with *measurement phase* as the within-subjects factor, and *treatment* as the between-subjects factor. This found a significant main effect of *phase,* whilst the remaining main effects and interactions were non-significant (after first log-transforming the data, *treatment:* $F_{1,14}$=0.279, *p*=0.605; *measurement phase:* $F_{1,14}$=4.906, *p*=0.044; *treatment*measurement phase:* $F_{1,14}$=0.115, *p*=0.739). Figure 3.5, which plots this data across *measurement phase,* by

treatment group, indicates that both *treatment* groups were faster to record a response in the third *measurement phase.*

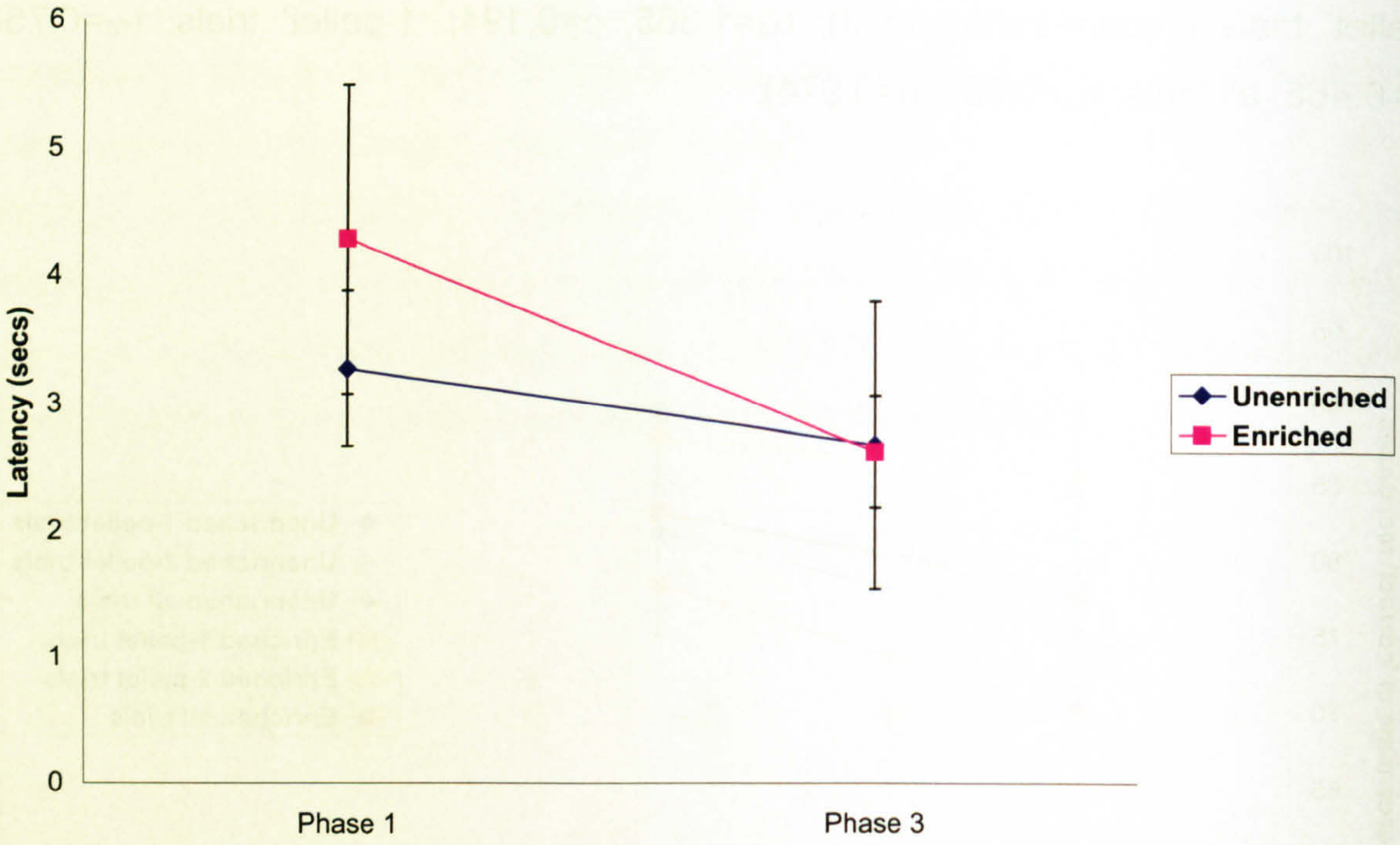

Figure 3.5 The mean latency to record *any* lever press response in the training sessions which took place on the consecutive days immediately prior to single-frequency probe testing, summarised by *measurement phase* (three such training sessions took place in each of the first, and third, *phases*), by *treatment* group (+/- 1SEM).

The latency to record a lever press response in only those trials associated with 2 pellets of food (i.e. the trials in which 2kHz was presented for the *2kHz=2pell contingency* group, and the trials in which 4kHz was presented for the *4kHz=2pell contingency* group) in the training sessions was also submitted to a repeated-measures GLM, with both *measurement phase* and *accuracy of response* (i.e. *'correct'* or *'incorrect'*) as within-subject factors, and *treatment* as a between-subjects factor. This found significant main effects of *measurement phase* and *accuracy of response*, whilst all other main effects and interactions were non-significant (after first log-transforming the data, *treatment:* $F_{1,14}=1.213$, $p=0.289$; *measurement phase:* $F_{1,14}=8.944$, $p=0.010$; *accuracy:* $F_{1,14}=33.265$, $p<0.001$; *measurement phase*treatment:* $F_{1,14}=0.214$, $p=0.650$; *accuracy*treatment:* $F_{1,14}=0.049$, $p=0.829$; *measurement phase*accuracy*: $F_{1,14}=0.015$, $p=0.904$;

*measurement phase*accuracy*treatment:* $F_{1,14}$=0.129, p=0.725). Figure 3.6, which plots this data by *accuracy* and *treatment,* across *measurement phase*, indicates that both *treatment* groups were quicker to make *correct* responses in both *measurement phases,* whilst responses were overall quicker in latter *phase* (as confirmed by inspection of the estimated marginal means derived from the model).

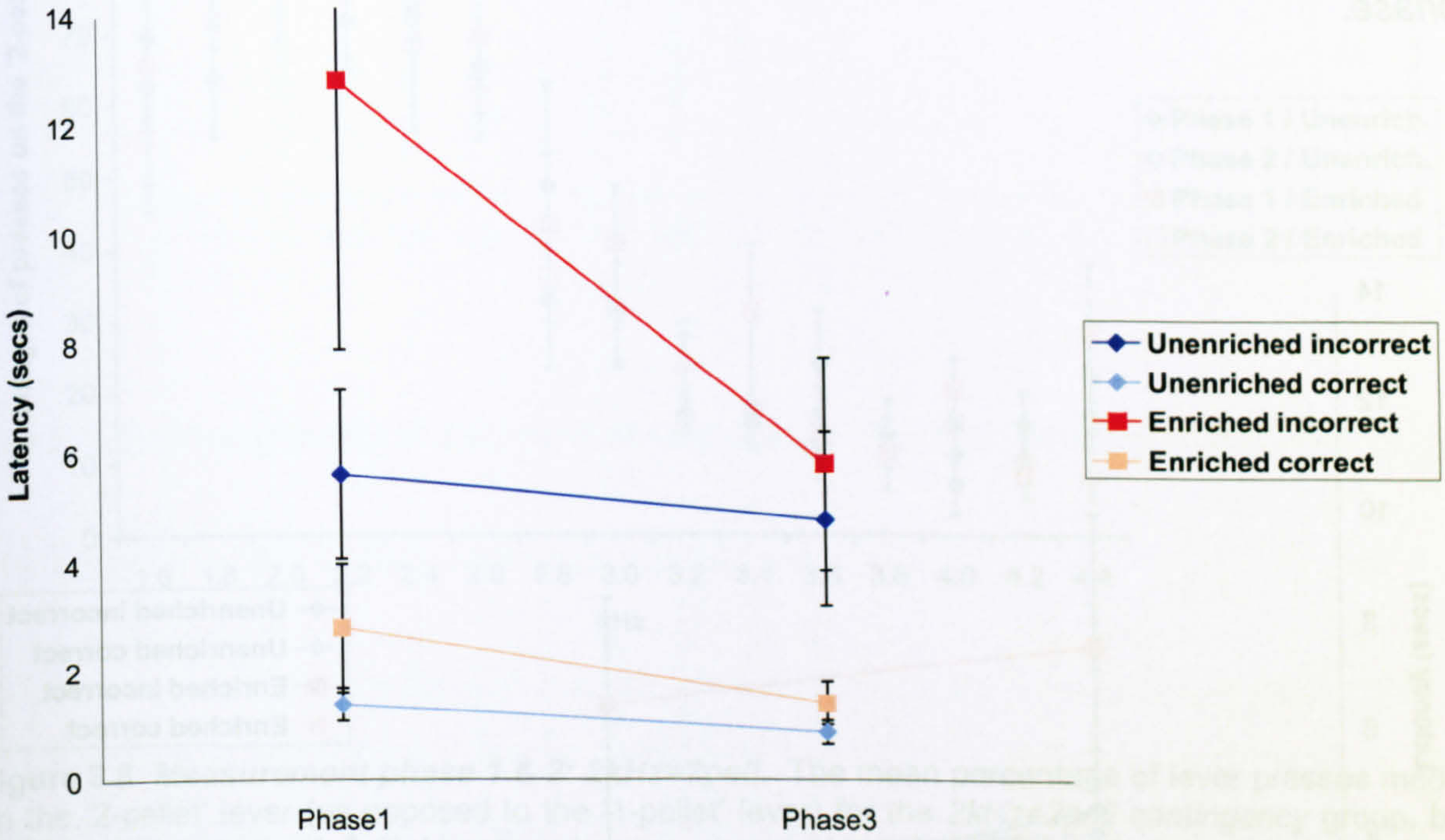

Figure 3.6 '2-pellet' trials only. The mean latency to record a 'correct' (i.e. press the '2-pellet' lever) or 'incorrect' (i.e. press the '1-pellet' lever) response in the trials associated with 2 pellets of food in the training sessions which took place on the consecutive days immediately prior to single-frequency probe testing, summarised by *measurement phase* (three such training sessions took place in each of the first, and third, *phases*), by *treatment* group (+/- 1SEM).

Finally, the latency to record a lever press response in only those trials associated with 1 pellet of food in the training sessions was also submitted to a repeated-measures GLM, of the same design as that employed above. The analysis found a significant main effect of *accuracy*, whilst all other main effects and interactions were non-significant (after first log-transforming the data, *treatment:* $F_{1,14}$=0.563, p=0.465; *measurement phase:* $F_{1,14}$=2.516, p=0.135; *accuracy:* $F_{1,14}$=5.910, p=0.029; *measurement phase*treatment:* $F_{1,14}$=0.488, p=0.496;

*accuracy*treatment:* $F_{1,14}$=1.743, *p*=0.208; *measurement phase*accuracy*: $F_{1,14}$=0.014, *p*=0.906; *measurement phase*accuracy*treatment:* $F_{1,14}$=0.094, *p*=0.763). An inspection of the estimated marginal means derived from the model indicated that the rats were overall faster to make 'correct' responses. Figure 3.7 plots this data by *accuracy* and *treatment,* across *measurement phase*; whilst it indicates that the *unenriched* group were quicker to make 'incorrect' responses (compared to 'correct' responses) in the last *phase,* this was largely attributable to a far outlying value, of particularly high latency, for 'correct' responding in the latter *phase.*

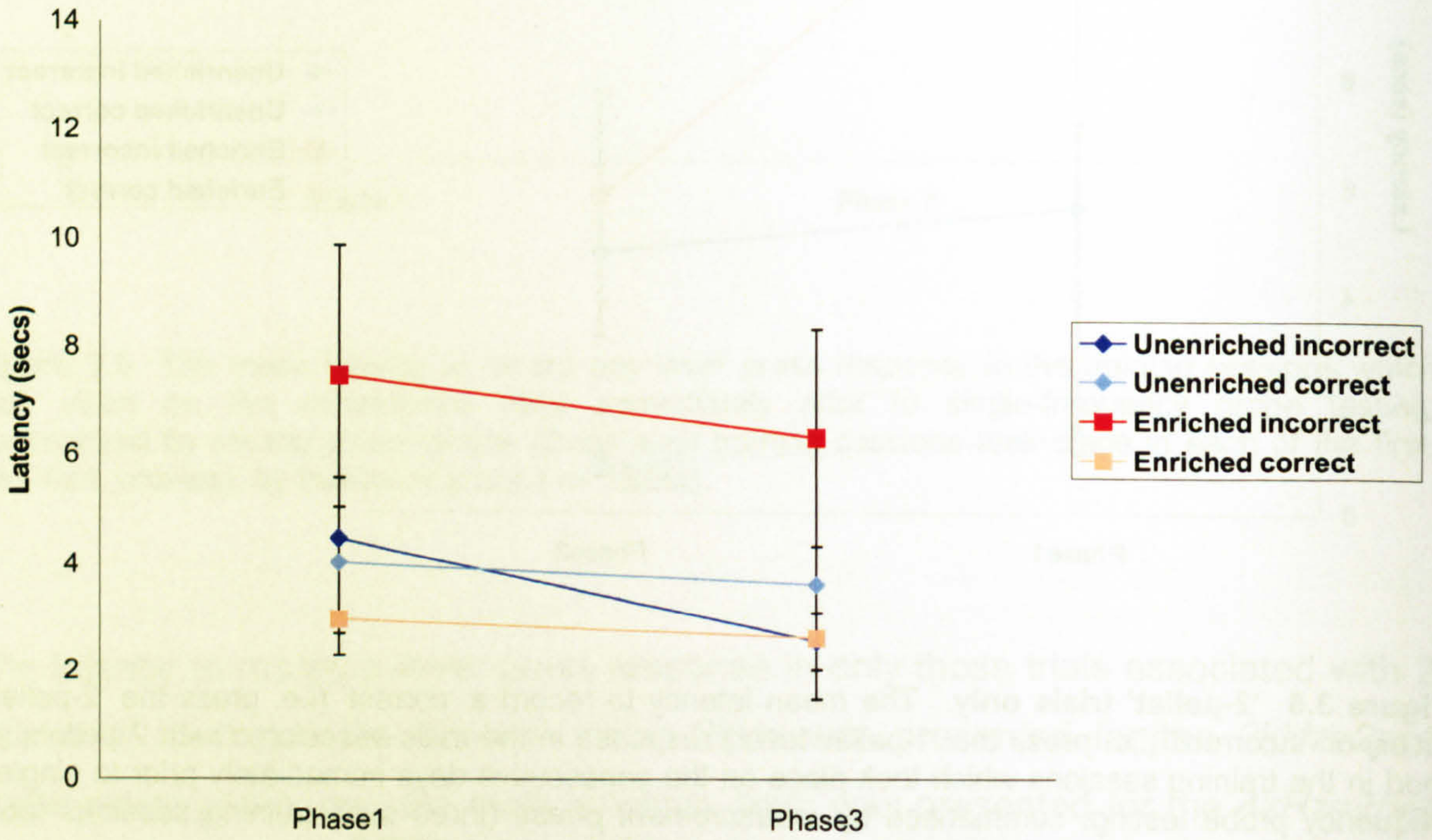

Figure 3.7 '1-pellet' trials only. The mean latency to record a 'correct' (i.e. press the '1-pellet' lever) or 'incorrect' (i.e. press the '2-pellet' lever) response in the trials associated with 1 pellets of food in the training sessions which took place on the consecutive days immediately prior to single-frequency probe testing, summarised by *measurement phase* (three such training sessions took place in each of the first, and third, *phases*), by *treatment* group (+/- 1SEM).

Single-frequency test sessions: lever choice

Figure 3.8 to Figure 3.11 chart the mean percentage of presses on the lever associated with 2 pellets of food (as opposed to the lever associated with 1 pellet

of food) across kHz, for the *2kHz=2pell contingency* group (comparing *measurement phase 1 & 2* (Figure 3.8), and *1 & 3* (Figure 3.9), respectively) and for the *4kHz=2pell contingency* group (comparing *measurement phase 1 & 2* (Figure 3.10), and *1 & 3* (Figure 3.11), respectively).

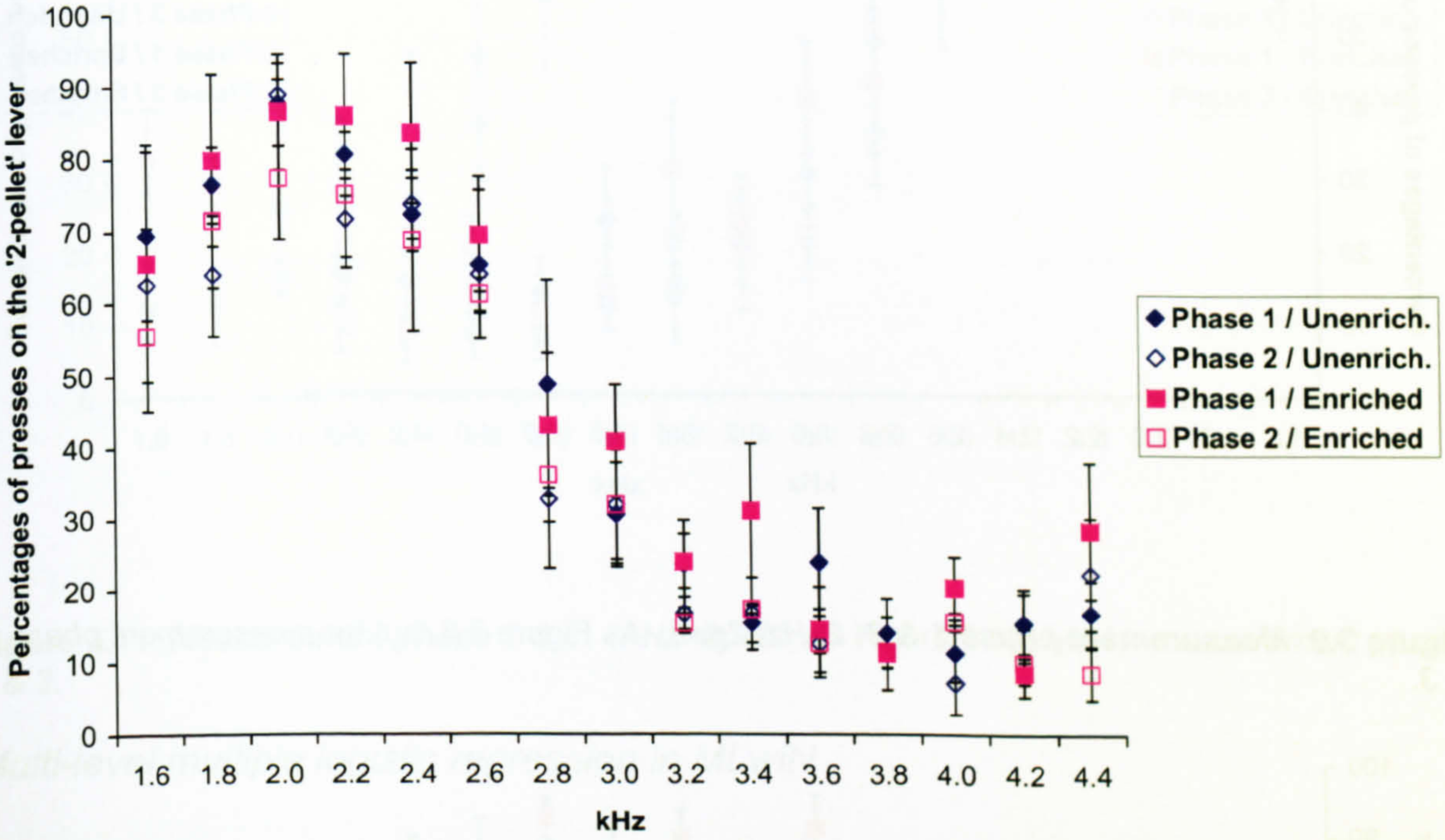

Figure 3.8 ***Measurement phase 1 & 2: 2kHz=2pell.*** The mean percentage of lever presses made on the '2-pellet' lever (as opposed to the '1-pellet' lever) for the *2kHz=2pell contingency* group, by *measurement phase (1 & 2) / treatment* group, across kHz (+/- 1 SEM. N.B. the means and SEM are derived from data summarised at the *subject*-level; in addition, the data pertaining to the 'reference tones' (i.e. 2kHz and 4kHz) are taken from the non-reinforced trials only).

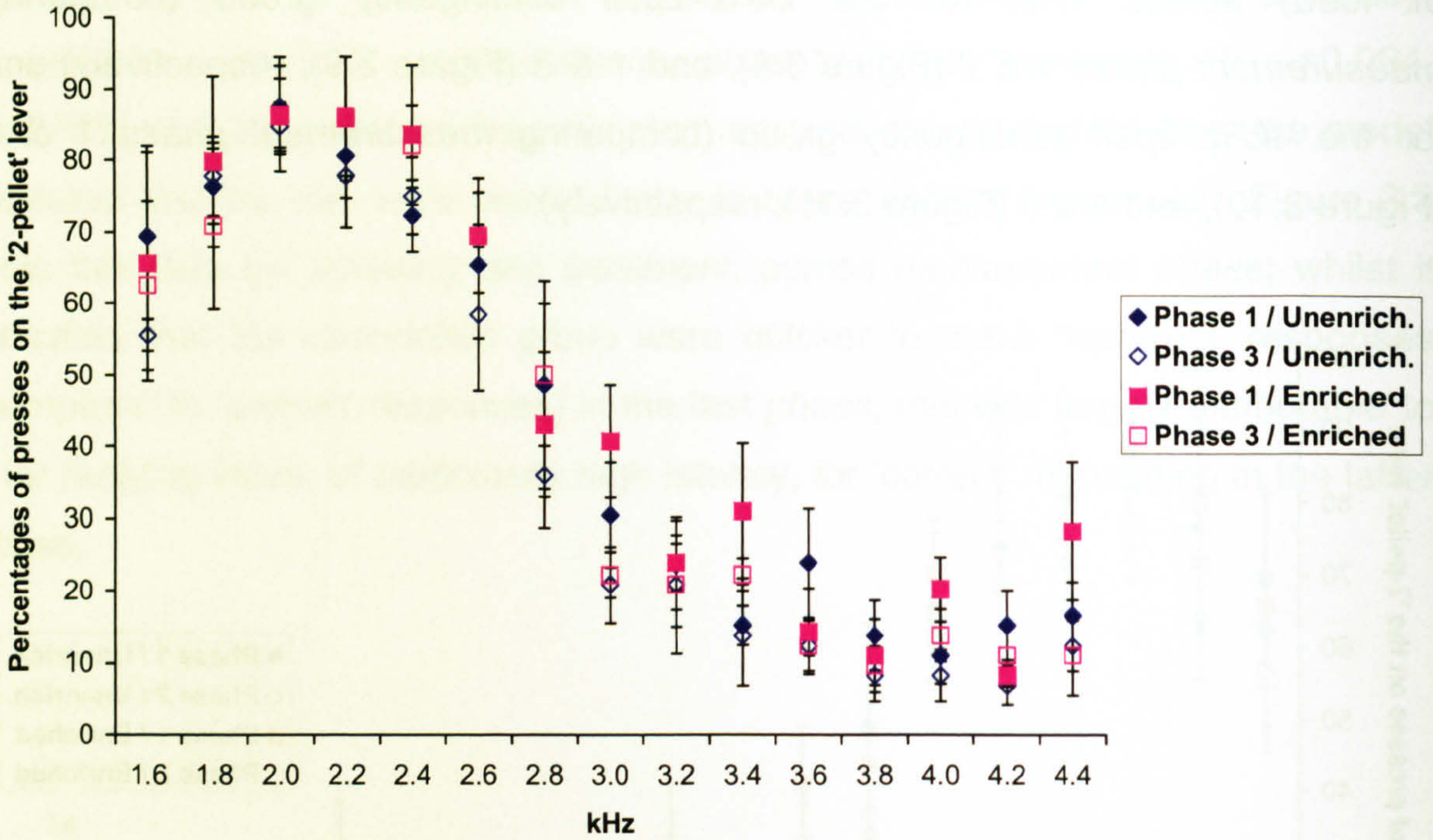

Figure 3.9 ***Measurement phase 1 & 3: 2kHz=2pell.*** As Figure 3.8, but for *measurement phase 1 & 3*.

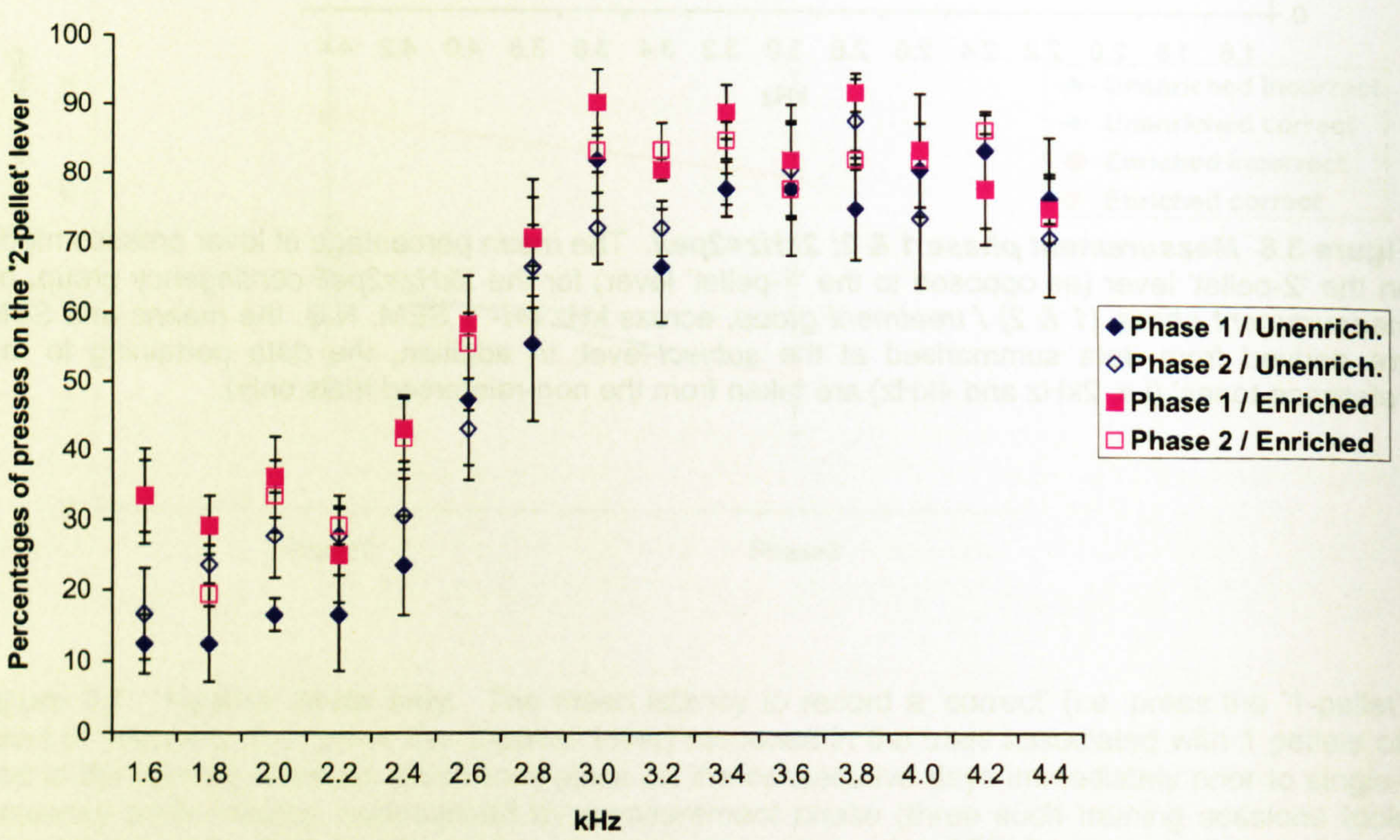

Figure 3.10 ***Measurement phase 1 & 2: 4kHz=2pell.*** As Figure 3.8, but for the *4kHz=2pell contingency* group.

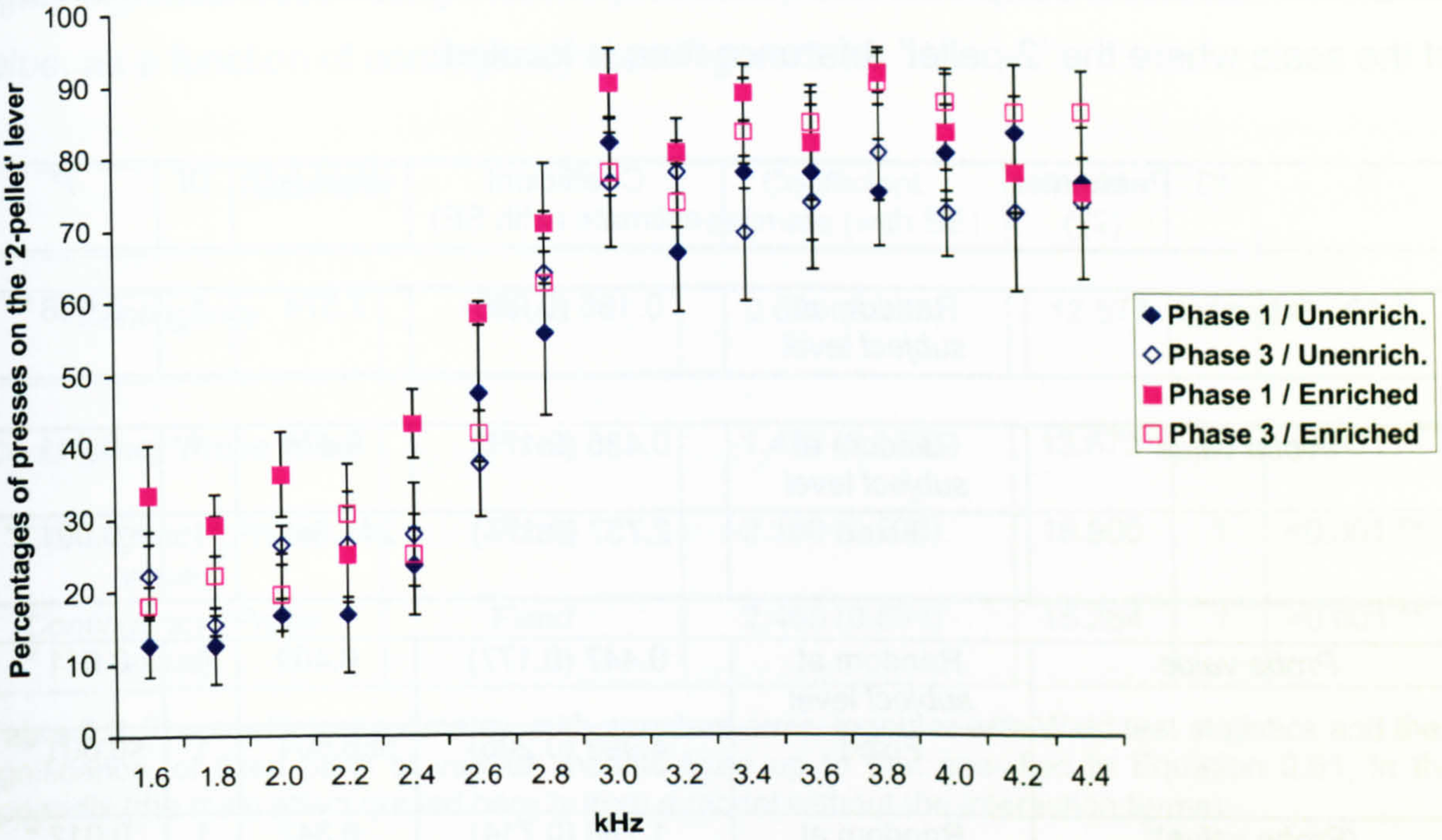

Figure 3.11 ***Measurement phase 1 & 3: 4kHz=2pell.*** As Figure 3.10, but for *measurement phase 1 & 3*.

Multi-level multiple logistic regression in MLwiN

As in the equivalent analysis in the last chapter (e.g. p.89), the dataset was defined as having two hierarchical levels, with *trial* (n=12,915) at the lowest level of the hierarchy (Level 1), and *subject* (n=16) at the next, higher, level of the hierarchy (Level 2; i.e. *trial* was nested within *subject*).

A 'random slope' model was again fitted, up to a cubic term for *probe value*[111], with the intercept, and the linear and quadratic *probe value* terms, allowed to vary at the *subject*-level. As Table 3.1 illustrates, this process revealed a significant difference between *subjects* both in the overall probability of pressing the '2-pellet' lever, and in the probability of doing so across *probe value*. There was also a significant fixed (overall) effect of *probe value*, indicating that the polynomial model was a good fit to the data, and that there was a significantly greater probability of pressing the '2-

[111] The standardised scale for *probe value*, as described on p.71

pellet' lever as the linear *probe value* term increased: i.e. as it approached the end of the scale where the '2-pellet' reference tone is located.

Parameter		Coefficient estimate (with SE)	Wald (χ^2)	Df	*P*
Intercept	Random at *subject* level	0.186 (0.068)	7.574	1	0.006 **
Probe value	Random at *subject* level	0.435 (0.171)	6.476	1	0.011 *
	Fixed	2.737 (0.174)	247.842	1	<0.001 **
Probe value	Random at *subject* level	0.447 (0.177)	6.409	1	0.011 *
	Fixed	4.992 (0.206)	586.504	1	<0.001 **
(Probe value)^2	Random at *subject* level	1.798 (0.714)	6.342	1	0.012 *
	Fixed	-0.979 (0.355)	7.602	1	0.006 **
(Probe value)^3	Fixed	-6.339 (0.296)	457.321	1	<0.001 **

Table 3.1 The coefficient estimates, with standard error, together with Wald test statistics and their significance, of fixed and random parts of various models fitted in gradual increments up to a polynomial 'random slope' model (as specified in Equation 0.50, in the Appendix).

The categorical predictor variables of main interest (*contingency, measurement phase* and *treatment*), were then systematically added to the model as fixed effects.

Contingency was first added to the model, initially as a main effect[112]. As Table 3.2 shows, as in the equivalent analysis in the last chapter, the *contingency* group *4kHz=2pell* had a significantly higher overall probability of pressing the lever associated with 2 pellets of food. The effect of *contingency* across *probe value* was then examined, with two-way interaction terms featuring *contingency* and each of the *probe value* predictors. As Table 3.2 indicates, these interactions were

[112] The *2kHz=2pell* group was the reference category, assigned a value of *'0'*, whilst the *4kHz=2pell* group was assigned a value of *'1'*.

highly significant, indicating that the probability of *lever choice* differs across *probe value,* as a function of *contingency:* see Figure 3.12.

Parameter		Coefficient estimate (with SE)	Wald (χ2)	Df	*P*
Contingency	Fixed	0.620 (0.175)	12.577	1	<0.001 **
*Contingency*Probe value*	Fixed	-1.436 (0.388)	13.672	1	<0.001 **
Contingency(Probe value)*^{2}	Fixed	-2.160 (0.526)	16.900	1	<0.001 **
Contingency(Probe value)*^{3}	Fixed	2.465 (0.631)	15.264	1	<0.001 **

Table 3.2 The coefficient estimates, with standard error, together with Wald test statistics and their significance, of fixed parts of various models fitted up to that specified in Equation 0.51, in the Appendix (the main effect quoted here is from a model without the interaction terms).

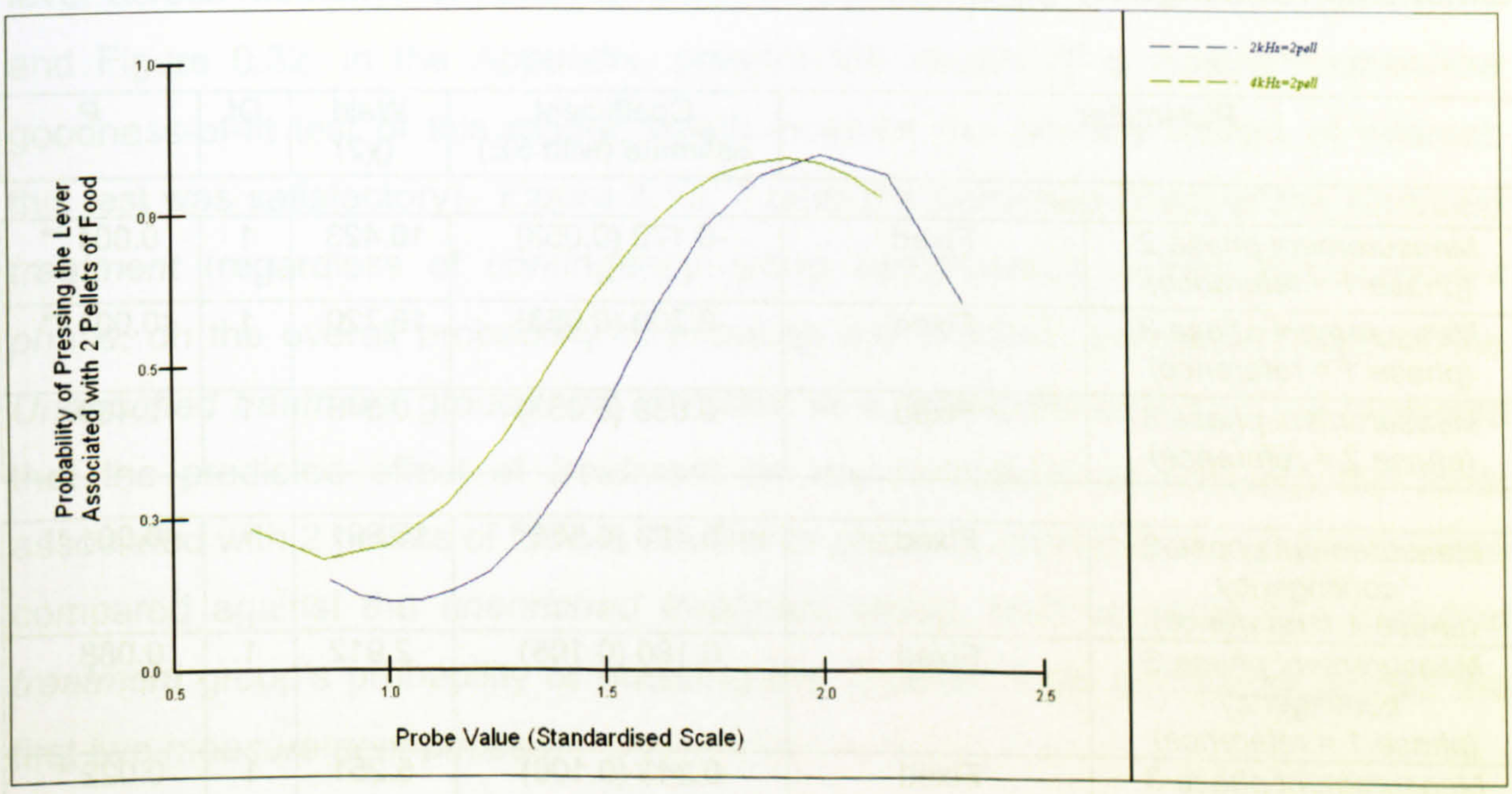

Figure 3.12 The predicted probability of pressing the '2-pellet' lever (as opposed to the '1-pellet' lever) by *contingency* group, across *probe value* (these predictions were generated from the model specified in Equation 0.51, in the Appendix).

Measurement phase was then added to the model as a main effect. As Table 3.3 indicates, there was a significantly lower probability of pressing the lever

associated with 2 pellets of food in both *Phase 2* and *Phase 3*, compared to *Phase 1*, but no significant difference between *Phase 2* and *Phase 3*. Therefore, if the model were to remain at this level of specification, there would be grounds for collapsing the latter two *measurement phases* into one, thus simplifying the model (e.g. Rasbash et al., 2005), but since we wish to further develop the model with additional terms, we will keep them as they are.

Various interactions of *measurement phase* with the *contingency* and *probe value* terms were then investigated: the two-way interaction between *measurement phase* and *contingency* was significant across all three *measurement phases*, except between *phase 1* and *phase 3* (see Table 3.3, and, for reference, Equation 0.53 & Figure 0.31, in the Appendix), but none of the interactions featuring *probe value* made a useful contribution to the model, indicating there was no change in the shape of the *lever choice* response curve in the various *measurement phases*, either when *contingency* group was taken into account, or when it was not.

Parameter		Coefficient estimate (with SE)	Wald (χ2)	Df	*P*
Measurement phase 2 (phase 1 = reference)	Fixed	-0.170 (0.053)	10.423	1	0.001 **
Measurement phase 3 (phase 1 = reference)	Fixed	-0.209 (0.053)	15.720	1	<0.001 **
Measurement phase 3 (phase 2 = reference)	Fixed	-0.039 (0.053)	0.543	1	0.461
*Measurement phase 2 *contingency (phase 1 = reference)*	Fixed	0.423 (0.106)	15.891	1	<0.001 **
*Measurement phase 3 *contingency (phase 1 = reference)*	Fixed	0.180 (0.105)	2.912	1	0.088
*Measurement phase 3 *contingency (phase 2 = reference)*	Fixed	-0.243 (0.106)	5.251	1	0.022 *

Table 3.3 The coefficient estimates, with standard error, of selected fixed parts of the model, together with Wald test statistics and their significance (the main effect is derived from the model specified in Equation 0.52, and the interaction term is derived from Equation 0.53 (both in the Appendix), each with appropriate alterations to the assignment of reference category).

Treatment was then added to the model as a main effect.[113] Table 3.4 indicates that whilst the *Enriched* group had a higher probability of pressing the lever associated with 2 pellets of food, this main effect did not explain a significant amount of the variance.

Various interactions of *treatment* with *measurement phase, contingency* and *probe value* were then explored, in a variety of models. There was no indication of any meaningful effect of *treatment* on the shape of the response curves (i.e. across *probe value*), either when other factors were taken into account (i.e. in interactions) or not, nor did the three-way interaction between the categorical predictors (i.e. *treatment, measurement phase* and *contingency)* explain a significant amount of variance. However, as Table 3.4 indicates, one of the key interactions (for the purposes of our hypotheses), between *treatment* and *measurement phase,* indicated a significant effect of *treatment* on the probability of pressing the '2-pellet' lever across *measurement phase,* when comparing *phase 1* to *phase 2* (Table 0.5 and Figure 0.32, in the Appendix, present the results of a Hosmer-Lemeshow goodness-of-fit test of this model, which includes the primary factors of interest: this test was satisfactory). Figure 3.13[114] plots the predicted effect of the *Enriched treatment* (regardless of *contingency* group assignment), across *measurement phase,* on the overall probability of pressing the '2-pellet', contrasted against the *Unenriched treatment* group, held constant at a probability of 0.5 [115]. It indicates that the predicted effect of *treatment* on the probability of pressing the lever associated with 2 pellets of food is smaller in *phase 2,* compared to *phase 1:* when compared against the *unenriched treatment* group, held constant, the *enriched treatment* group's probability of pressing the '2-pellet' lever decreases across the first two *measurement phases.*

[113] The *Unenriched treatment* group were nominated as the reference category, with a value of *'0'*, and the *Enriched treatment* group were assigned a value of *'1'.*

[114] See Figure 0.33, in the Appendix, for a similar chart, this time allowing movement in the prediction curve pertaining to the *unenriched* group as well. It is derived from a slightly different equation to that used to generate the corresponding chart presented in the main text (necessarily so), but whilst the equation has a slightly different range of predictors, the relevant coefficient estimates are very similar, and so the resulting chart is unlikely to be misleading.

[115] The equation used to generate this plot is the difference between the prediction equation for the *enriched* group (the whole of the equation specified in Equation 0.55, minus *subject*-level random effects), and the prediction equation for the *unenriched group* (the whole of that equation minus all the terms featuring *treatment,* since the *unenriched* group is assigned the reference category, with a value of *'0'):* i.e. *treatment + treatment*measurement phase* (e.g. p.73 of Rasbash et al., 2005).

Parameter		Coefficient estimate (with SE)	Wald ($\chi 2$)	Df	*P*
Treatment	Fixed	0.274 (0.161)	2.886	1	0.089
*Treatment*Measurement phase 2 (phase 1 = reference)*	Fixed	-0.284 (0.106)	7.213	1	0.007 **
*Treatment*Measurement phase 3 (phase 1 = reference)*	Fixed	-0.144 (0.105)	1.862	1	0.172
*Treatment*Measurement phase 3 (phase 2 = reference)*	Fixed	0.140 (0.105)	1.768	1	0.184

Table 3.4 The coefficient estimates, with standard error, of selected fixed parts of the model, together with Wald test statistics and their significance (the main effect is derived from the model specified in Equation 0.54, in the Appendix, and the interaction term is derived from Equation 0.55, also in the Appendix, with appropriate alterations to the assignment of reference category).

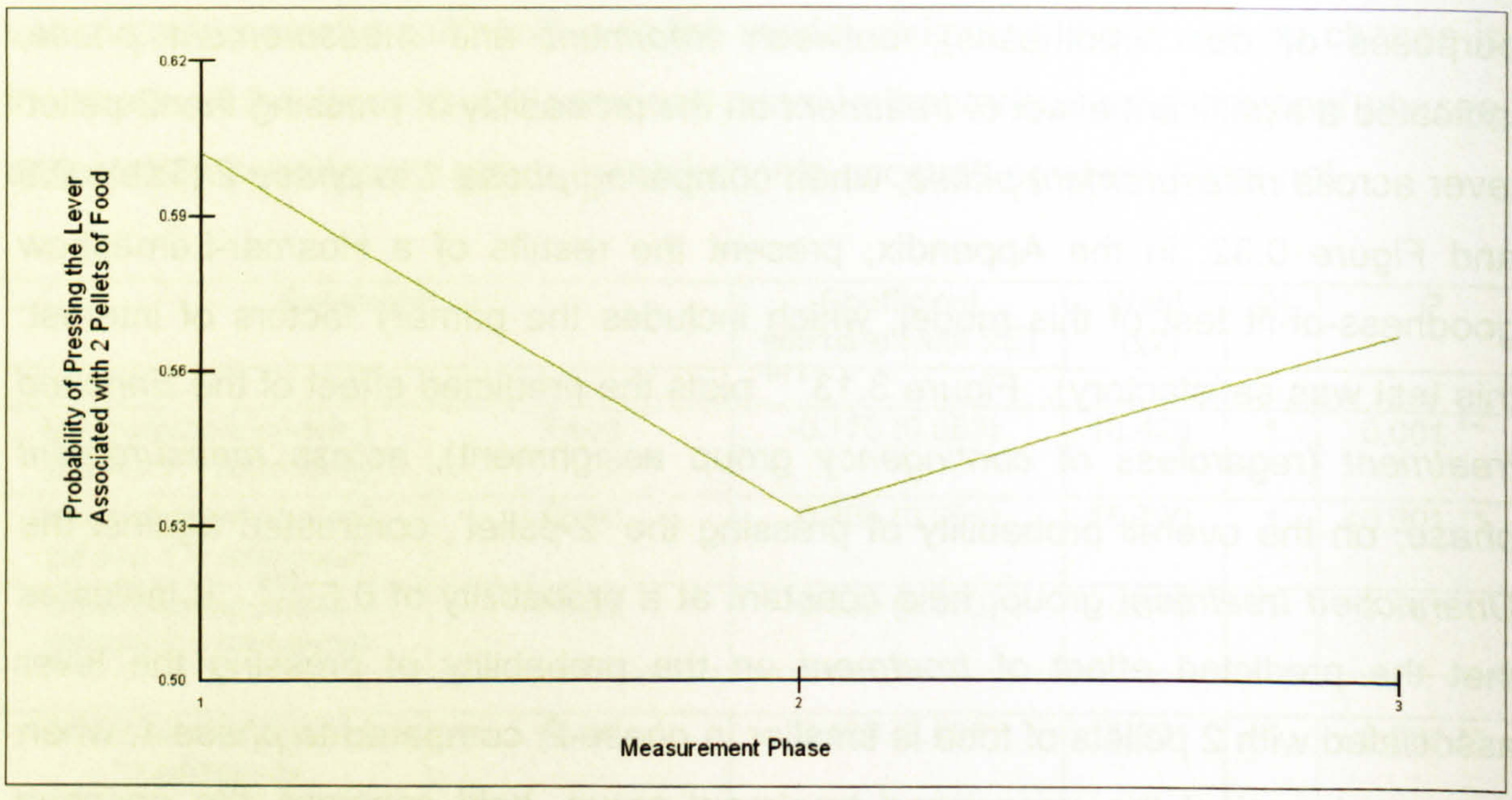

Figure 3.13 The predicted effect of being in the *enriched treatment* group on the probability of pressing the '2-pellet' lever, across different levels of *measurement phase* (from a prediction equation derived from the model specified in Equation 0.55). Here, the *enriched treatment* group is compared against the reference category of the *unenriched treatment* group, which has a probability held constant at 0.5 (if the *enriched treatment* group was assigned the reference category, and the *unenriched treatment* group contrasted against that, then the resulting prediction plot would resemble a mirror image of the above, with the mirror held along the x-axis; for an alternative way of plotting such data, see Figure 0.33, in the Appendix).

Finally, a number of other terms of interest were added to the model. As Table 3.5 indicates, the addition of *Prior treatment* (i.e. the *treatment* employed in the previous study in which the rats were subject, described in Chapter 2) as a main

effect[116], found that *subjects* previously in the *UHT (unpredictable housing)* group were significantly more likely to press the lever associated with 2 pellets of food compared to those *subjects* previously in the *control (predictable housing)* group.

A two-way interaction between *prior treatment* (i.e. *UHT* or *control*) and *current treatment* (i.e. *enriched* or *unenriched*) was added to the model, to investigate whether treatment history interacted with the current treatment regime: as Table 3.5 also indicates, however, this was not a significant term. We also investigated any change in the effect of *prior treatment* on the probability of pressing the '2-pellet' lever across *measurement phase,* by adding an interaction between these terms. As Table 3.5 indicates, this interaction was significant when comparing the first and the third *measurement phase.* Figure 3.14[117] plots the predicted effect of the *unpredictable housing treatment* on the overall probability of pressing the '2-pellet' lever in the current experiment, across *measurement phase*: the main effect of *prior treatment* is apparent, with those *subjects* previously in the *UHT* group consistently more likely to press the '2-pellet' lever than the *control (predictable housing)* group (held constant, with a probability of 0.5), but the chart illustrates this difference diminishes as *measurement phase* progresses, with the change between the first and third *measurement phases* being the greatest.

As Table 3.5 additionally indicates, interactions of *prior treatment* with the *probe value* terms were also significant, indicating there was a change in the probability of *lever choice*, across *probe value,* depending on whether *subjects* were previously in the *UHT* group or not. Figure 3.15, which plots the resulting predictions, again illustrates the main effect of *prior treatment,* with the *UHT* group overall more likely to press the '2-pellet' lever compared to the *control (predictable housing)* group[118], but further indicates that this difference is particularly pronounced around the '2-pellet' reference tone, and around the *probe values* at the far end of the scale, beyond the '1-pellet' reference tone.

---

[116] The *control (predictable housing)* group were assigned the reference category, with a value of *'0'*, and the *unpredictable housing treatment (UHT)* group were assigned a value of *'1'*.

[117] As before, see Figure 0.34, in the Appendix, for an alternative method of plotting such predictions, again derived from a slightly different equation.

[118] Again, held constant, with a probabilty of 0.5.

Parameter		Coefficient estimate (with SE)	Wald ($\chi 2$)	Df	*P*
Prior treatment	Fixed	0.324 (0.141)	5.263	1	0.022 *
*Prior treatment*Treatment*	Fixed	0.048 (0.282)	0.028	1	0.867
*Prior treatment *Measurement phase 2 (phase 1 = reference)*	Fixed	-0.055 (0.106)	0.265	1	0.607
*Prior treatment *Measurement phase 3 (phase 1 = reference)*	Fixed	-0.231 (0.105)	4.819	1	0.028 *
*Prior treatment *Measurement phase 3 (phase 2 = reference)*	Fixed	-0.177 (0.106)	1.793	1	0.181
*Prior treatment*Probe Value*	Fixed	0.871 (0.381)	5.225	1	0.022 *
Prior treatment(Probe Value)^2*	Fixed	0.578 (0.543)	1.133	1	0.287
Prior treatment(Probe Value)^3*	Fixed	-1.673 (0.598)	7.836	1	0.005 **

Table 3.5 The coefficient estimates, with standard error, of selected fixed parts of the model, together with Wald test statistics and their significance (derived from the models specified in Equation 0.57 to Equation 0.60 (model featuring non-significant *Prior treatment*treatment* term not shown), in the Appendix, with appropriate alterations to the assignment of reference category).

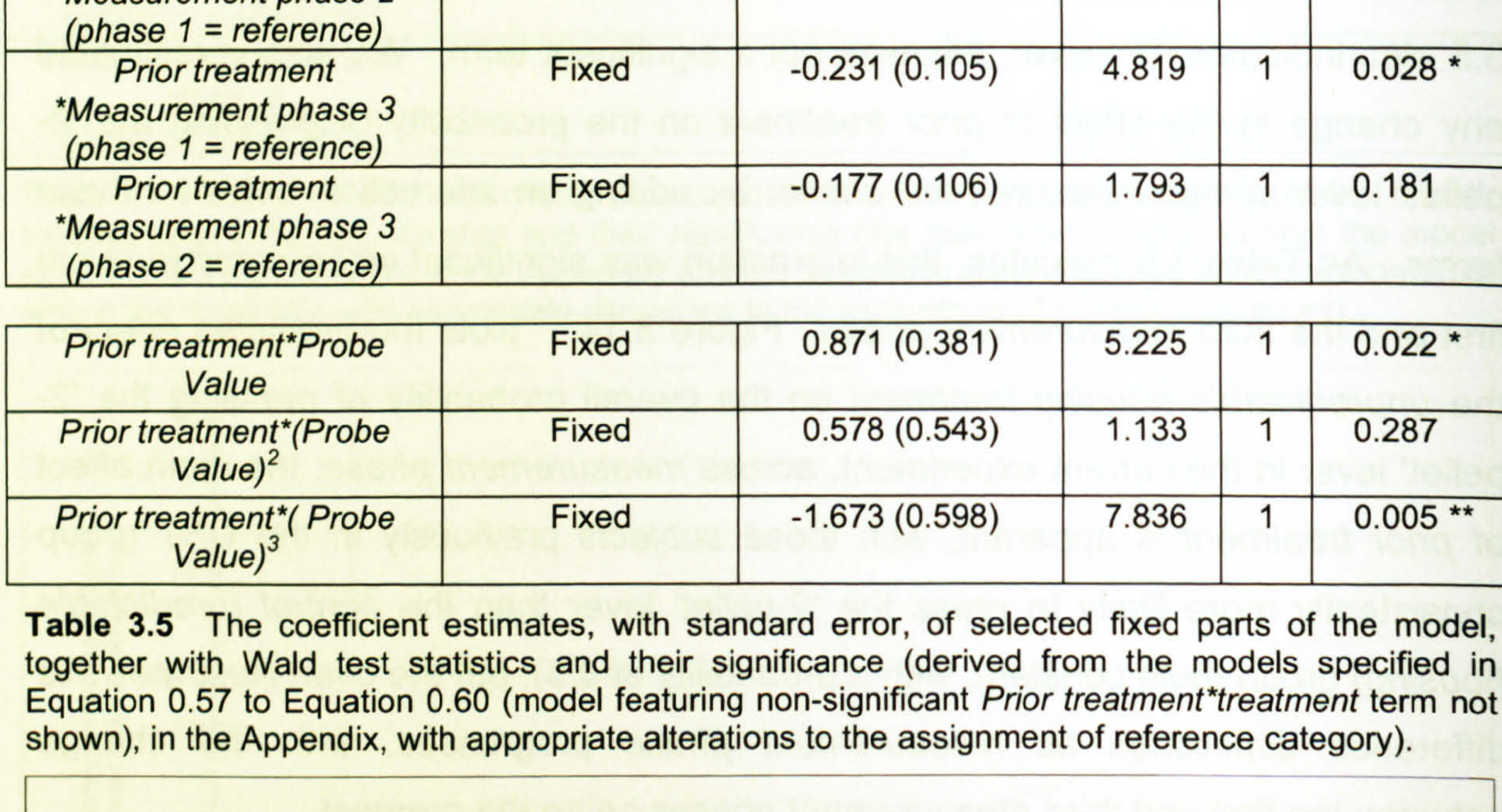

Figure 3.14 The predicted effect of being in the *unpredictable housing treatment (UHT)* group on the probability of pressing the '2-pellet' lever, across different levels of *measurement phase* (from a prediction equation derived from the model specified in Equation 0.58). Here, the *UHT* group is compared against the reference category of the *control (predictable housing)* group, which has a probability held constant at 0.5 (for an alternative way of plotting such data, see Figure 0.34, in the Appendix).

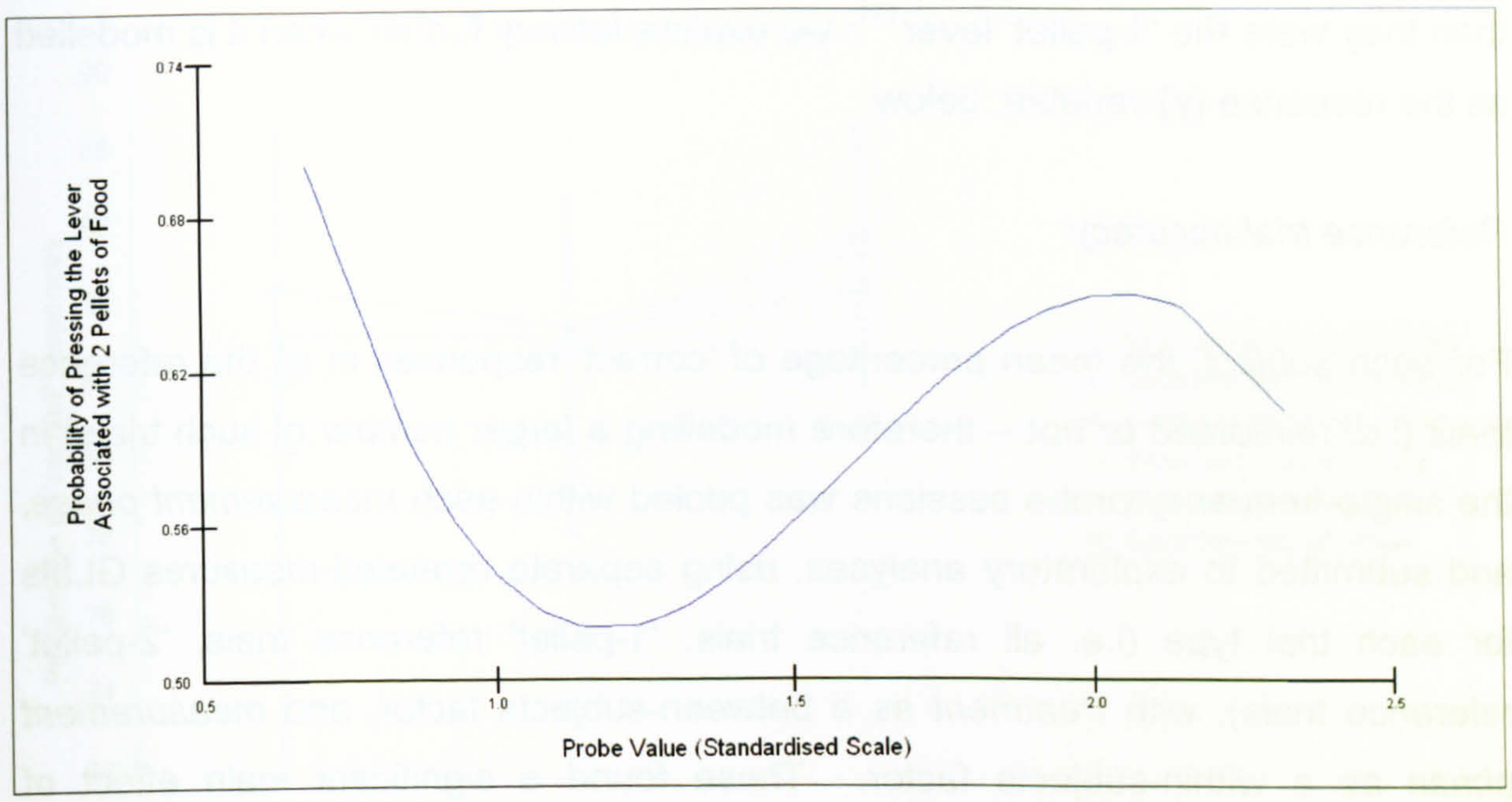

Figure 3.15 The predicted effect of being in the *unpredictable housing treatment (UHT)* group on the probability of pressing the '2-pellet' lever, across *probe value* (from a prediction equation derived from the model specified in Equation 0.60) Here, the *UHT* group is compared against the reference category of the *control (predictable housing)* group, which has a probability held constant at 0.5.

As in the analysis conducted on the unpredictable housing experimental data described in the last chapter, the *Position of the 2-pellet lever* was added to the model.[119] As Table 3.6 indicates, unlike that analysis, here the term did not explain a significant amount of the variance.[120]

Parameter		Coefficient estimate (with SE)	Wald (x2)	Df	*P*
Position of the 2-pellet lever	Fixed	0.118 (0.137)	0.745	1	0.388
Latency	Fixed	-0.020 (0.005)	15.192	1	<0.001 **

Table 3.6 The coefficient estimates, with standard error, of selected fixed parts of the model, together with Wald test statistics and their significance (from the models specified in Equation 0.62, and Equation 0.63, respectively: both in the Appendix).

Finally, *latency* to record a lever press was added to the model as a main effect. As Table 3.6 shows, the rats were significantly quicker to press the '2-pellet' lever

---

[119] The counterbalanced group which had the '2-pellet' lever to the left of the food hopper were assigned the reference category, with a value of *'0'*, whilst those with the '2-pellet' lever to the right of the food hopper were assigned a value of *'1'*.

[120] Nor did *Room* (i.e. the room in which the rat's were trained and tested: either Room *A* or *B*), nor the interaction between the two terms (not reported).

than they were the '1-pellet' lever[121]: we explore *latency* further when it is modelled as the response (y) variable, below.

Reference trial accuracy

For each *subject*, the mean percentage of 'correct' responses in *all* the reference trials (i.e. reinforced or not – therefore modelling a larger number of such trials) in the single-frequency probe sessions was pooled within each *measurement phase,* and submitted to exploratory analyses, using separate repeated-measures GLMs for each trial type (i.e. all reference trials, '1-pellet' reference trials, '2-pellet' reference trials), with *treatment* as a between-subjects factor, and *measurement phase* as a within-subjects factor. These found a significant main effect of *measurement phase*, and a significant interaction between *measurement phase* and *treatment*, when modelling accuracy across all reference trials, and when modelling accuracy in just the '1-pellet' reference trials; all other main effects and interactions were non-significant (all reference trials - *treatment:* $F_{1,14}$=0.056, p=0.817; *measurement phase:* $F_{1.9,26.602}$=8.184, p=0.002; *treatment*measurement phase*: $F_{2,26.602}$=3.502, p=0.047; '2-pellet' reference trials - *treatment:* $F_{1,14}$=0.384, p=0.545; *measurement phase:* $F_{1.731,24.238}$=0.678, p=0.497; *treatment*measurement phase*: $F_{1.731,24.238}$=0.284, p=0.724; '1-pellet' reference trials - *treatment:* $F_{1,14}$=0.760, p=0.398; *measurement phase:* $F_{1.958,27.410}$=4.633, p=0.019; *treatment*measurement phase*: $F_{1.958,27.410}$=3.685, p=0.039). Inspection of the (non-corrected) interaction contrasts indicated significance when comparing the first and third *measurement phases,* in both 'all reference' ($F_{1,14}$=7.954, p=0.014) and '1-pellet reference' ($F_{1,14}$=6.07, p=0.027) analyses, whilst the interaction contrast neared significance when comparing the first and second *measurement phase* in the '1-pellet reference' analysis ($F_{1,14}$=3.941, p=0.067); all other contrasts were non-significant. Figure 3.16 plots this data, summarised by reference trial type and *treatment*, across *measurement phase*; it indicates that the accuracy of the *enriched* group in the '1-pellet' reference trials substantially improved across *measurement phase,* as did their overall accuracy.

---

[121] This remained the case when only datapoints of a *latency* under, or equal to, 3 seconds (n=11,987) were modelled (p<0.001).

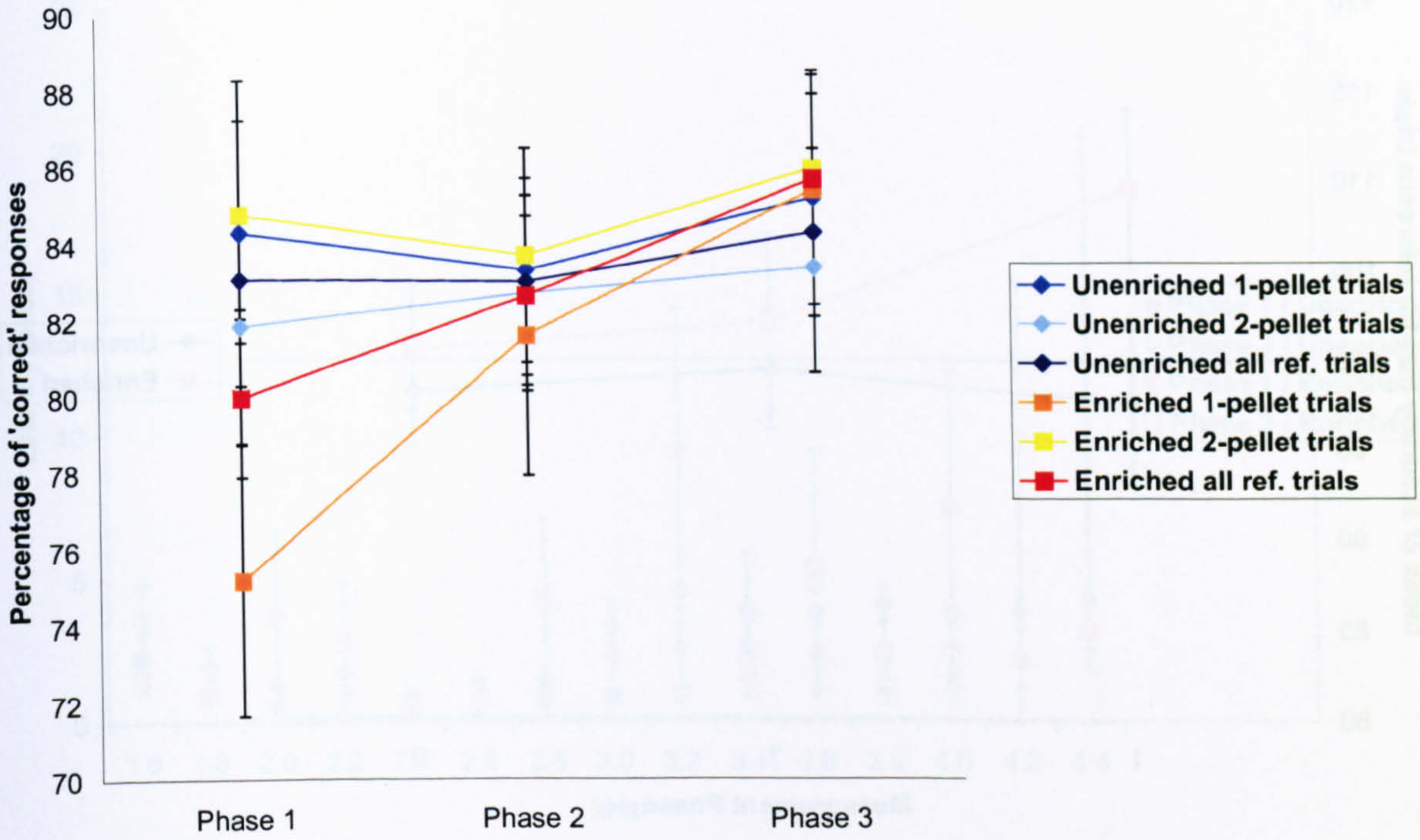

Figure 3.16 The mean percentage of 'correct' responses in the reference trials (both reinforced, and non-reinforced) in the single-frequency probe sessions, summarised by reference trial type, *measurement phase*, and *treatment* (+/- 1SEM).

For reference, Figure 3.17 plots this data in a slightly different manner: adding the percentage accuracy in the '2-pellet' reference trials to the percentage *inaccuracy* in the '1-pellet' reference trials; as such, an index of error bias in the reference trials is produced, with values above 100 indicating a bias towards pressing the '2-pellet' lever, and values below 100 indicating the opposite (i.e. a bias towards pressing the '1-pellet' lever). As such, it confirms the pattern suggested by the analyses above: namely that the *enriched* group had a bias towards pressing the '2-pellet' lever at 'baseline' (i.e. during *phase 1),* but this gradually attenuated to a point in the final *measurement phase* in which there was very little net bias (i.e. they were equally accurate, or close to it, in both types of reference trial).

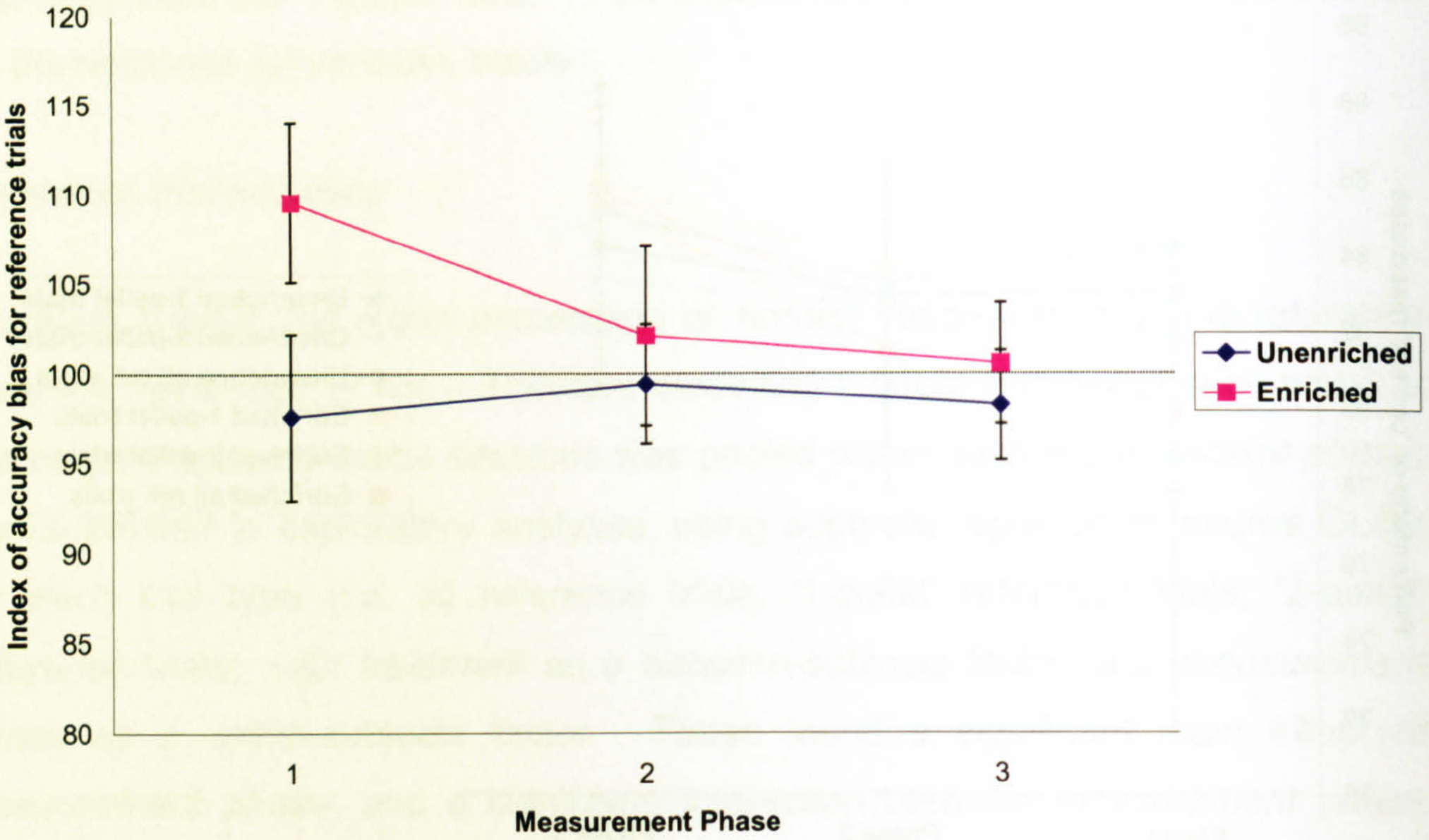

Figure 3.17 The mean index of accuracy bias (see main text) in the reference trials presented during the single-frequency probe sessions, summarised by *treatment* and *measurement phase* (+/- 1SEM); values above 100 indicate a bias towards pressing the '2-pellet' lever, whilst values below 100 indicate a bias towards pressing the '1-pellet' lever.

Single-frequency test sessions: latency

Figure 3.18 to Figure 3.21 plot the mean latency to press the lever associated with 2 pellets of food, across kHz, for the *2kHz=2pell* (Figure 3.18 compares *phase 1* and *phase 2;* Figure 3.19 compares *phase 1* and *phase 3*), and *4kHz=2pell* (Figure 3.20 compares *phase 1* and *phase 2;* Figure 3.21 compares *phase 1* and *phase 3*) *contingency* groups, respectively, whilst Figure 3.22 to Figure 3.25 plot the mean latency to press the lever associated with 1 pellet of food, across kHz, for the *2kHz=2pell* (Figure 3.22 compares *phase 1* and *phase 2;* Figure 3.23 compares *phase 1* and *phase 3*), and *4kHz=2pell* (Figure 3.24 compares *phase 1* and *phase 2;* Figure 3.25 compares *phase 1* and *phase 3*) *contingency* groups, respectively.

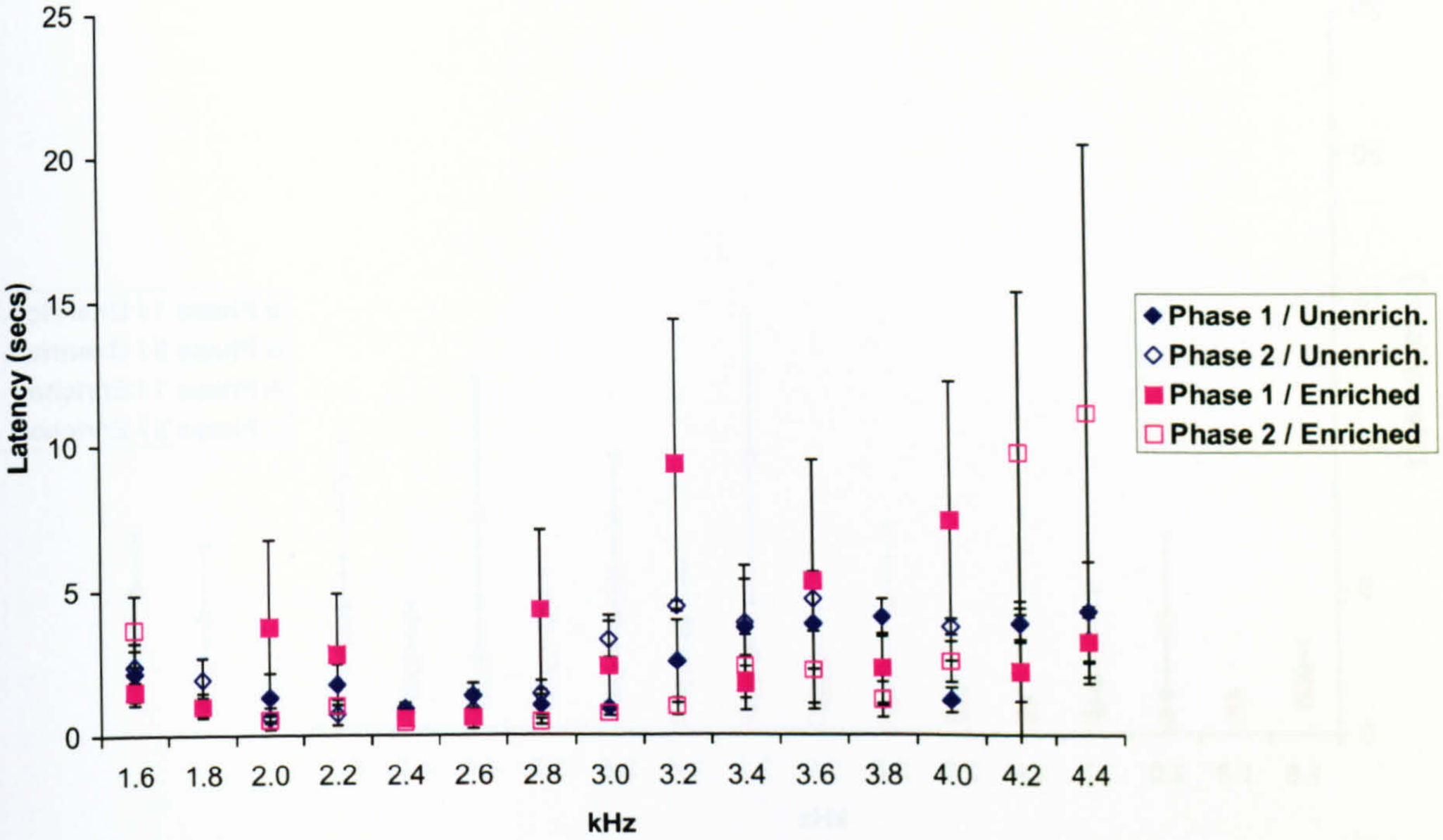

Figure 3.18 '2-pellet' lever latency: *2kHz=2pell / Measurement Phase 1 & 2.* The mean latency, in seconds, for the *2kHz=2pell contingency* group to press the lever associated with 2 pellets of food, by *treatment* group, across kHz, for the first two *measurement phases* (+/- 1 SEM. N.B. the means and SEM are derived from data summarised at the *subject*-level (note: if, in one of the *measurement phases*, a *subject* never pressed this particular lever following one of the kHz tones, that *subject* did not contribute to the corresponding summary datapoint[122]); in addition, the data pertaining to the 'reference tones' (i.e. 2kHz and 4kHz) are taken from the non-reinforced trials only).

[122] An alternative would have been to enter a maximum 'timed-out' latency, but since each trial terminated following a lever press (rather than being timed-out), it is not clear what value any such maximum latency would take.

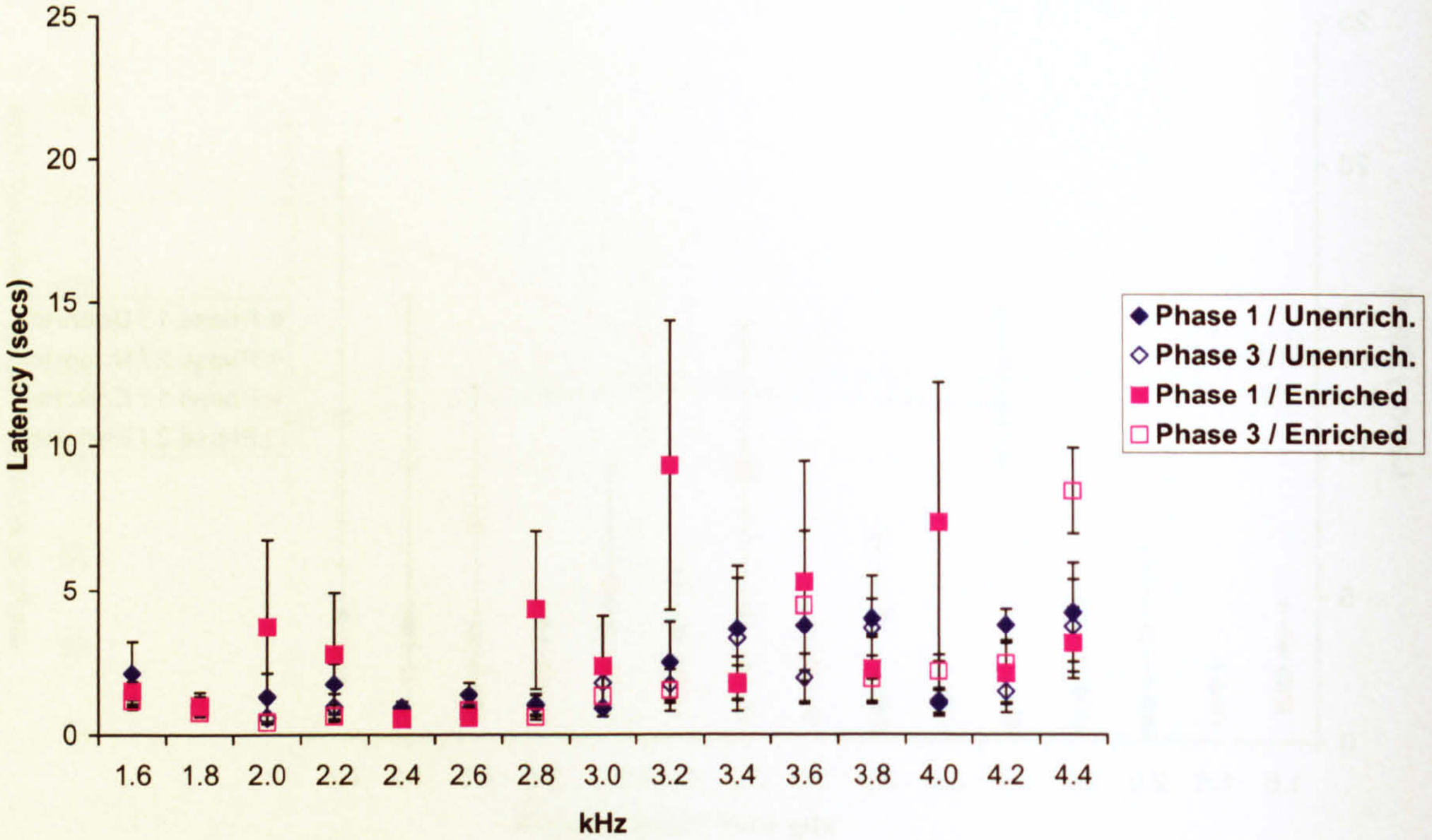

Figure 3.19 '2-pellet' lever latency: *2kHz=2pell / Measurement Phase 1 & 3.* As Figure 3.18, but plotting the first and last *measurement phase.*

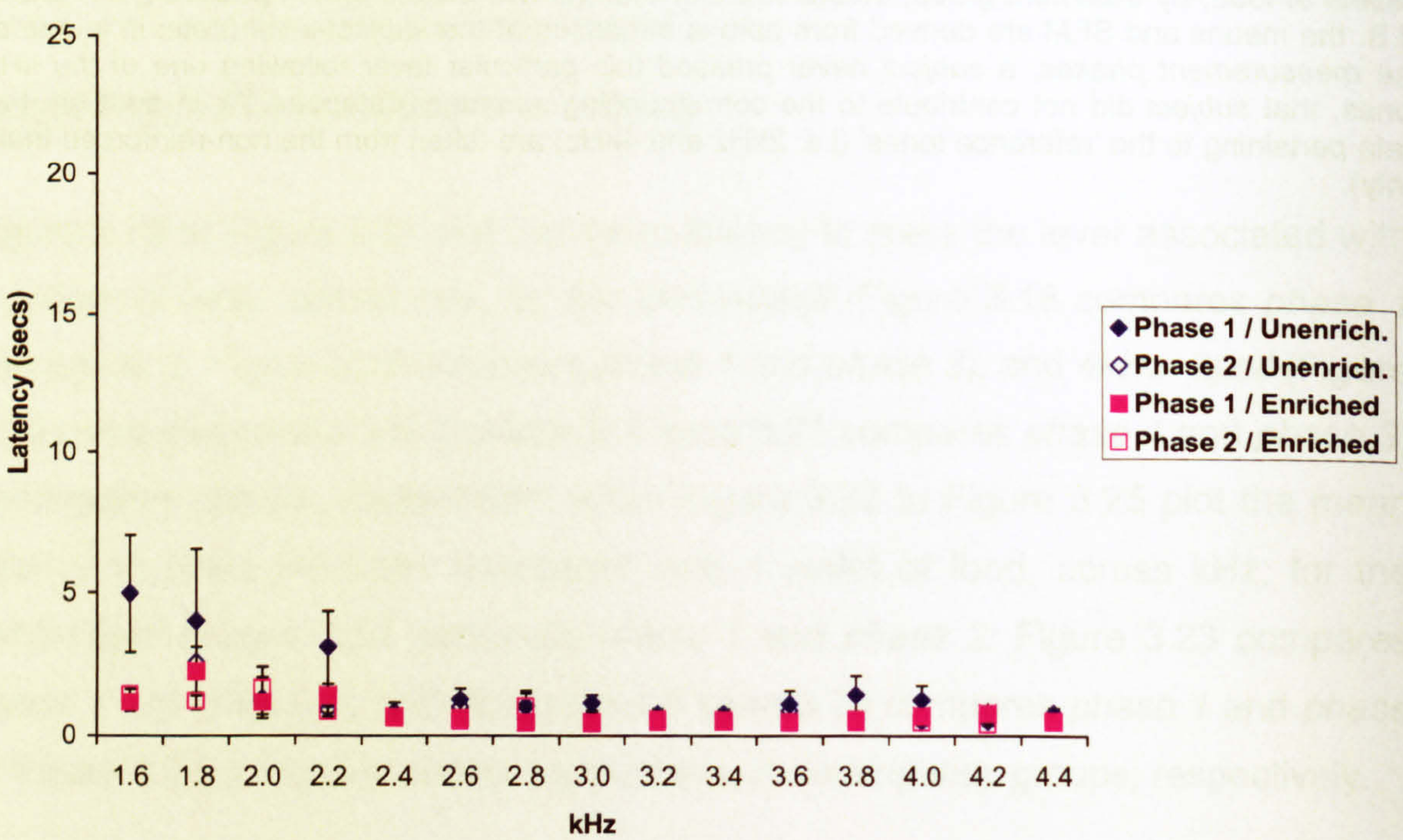

Figure 3.20 '2-pellet' lever latency: *4kHz=2pell / Measurement Phase 1 & 2.* As Figure 3.18, but for the *4kHz=2pell contingency* group.

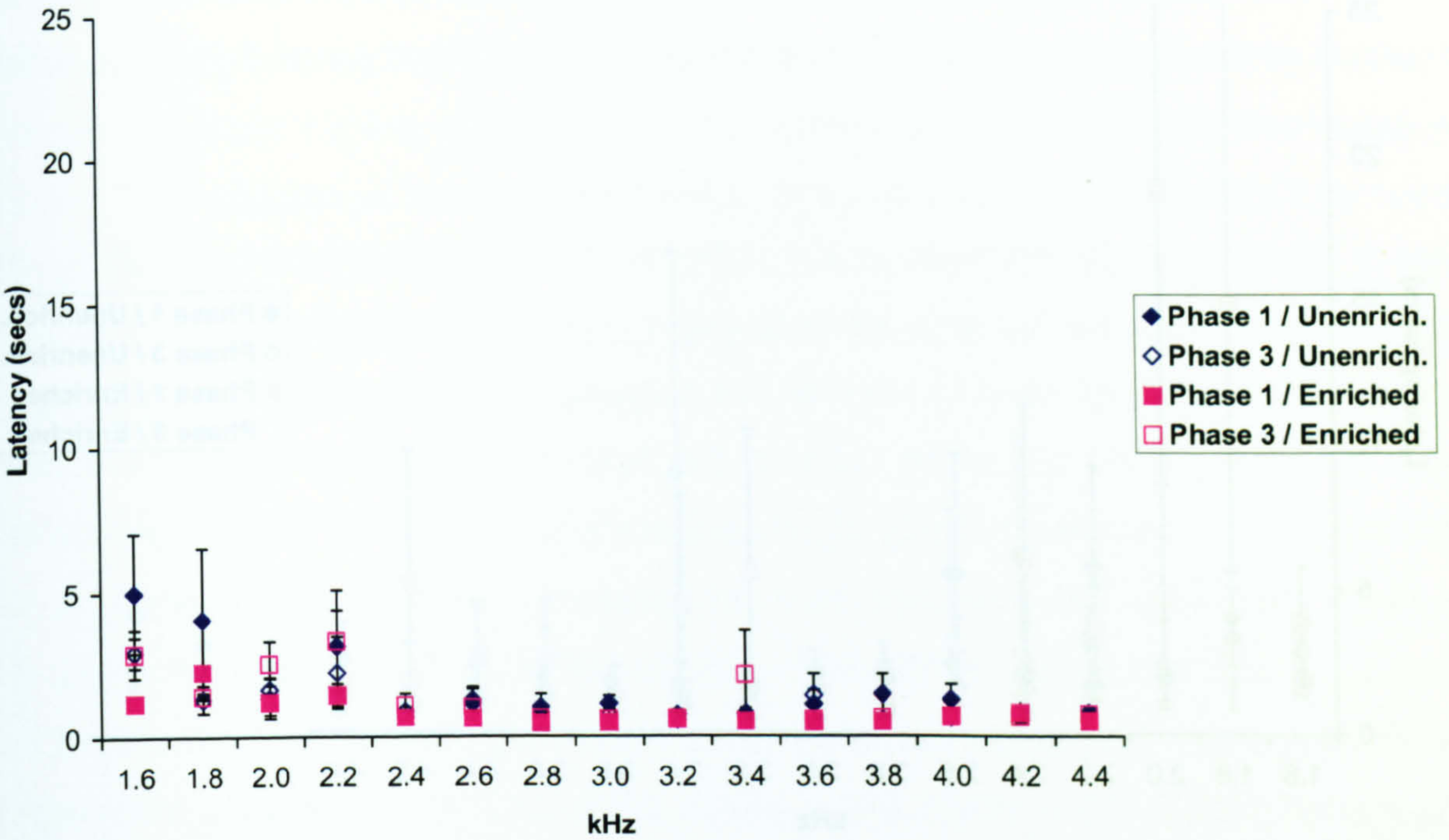

Figure 3.21 '2-pellet' lever latency: *4kHz=2pell / Measurement Phase 1 & 3.* As Figure 3.20, but plotting the first and last *measurement phase.*

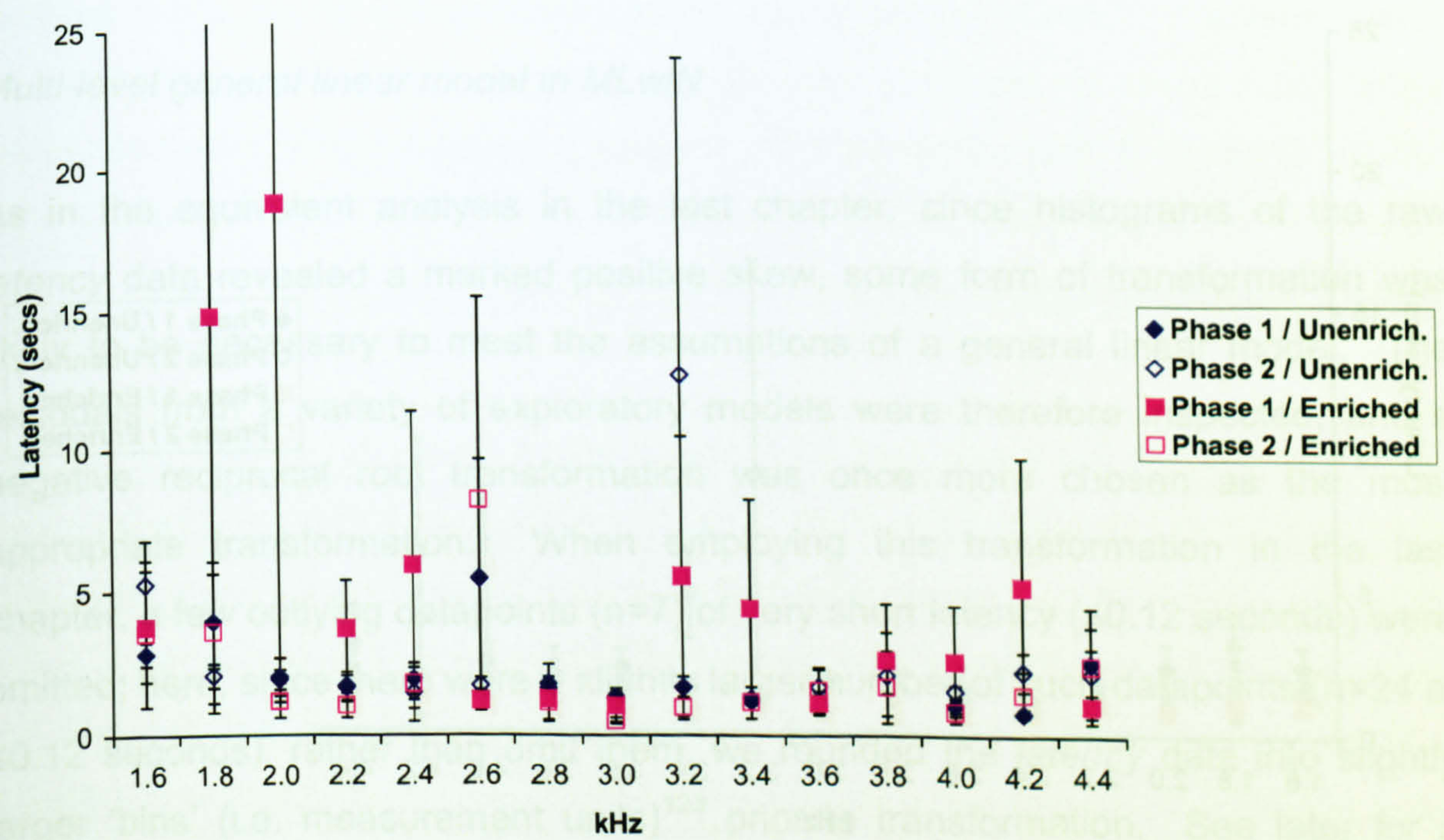

Figure 3.22 '1-pellet' lever latency: *2kHz=2pell / Measurement Phase 1 & 2.* As Figure 3.18, but for presses on the '1-pellet' lever.

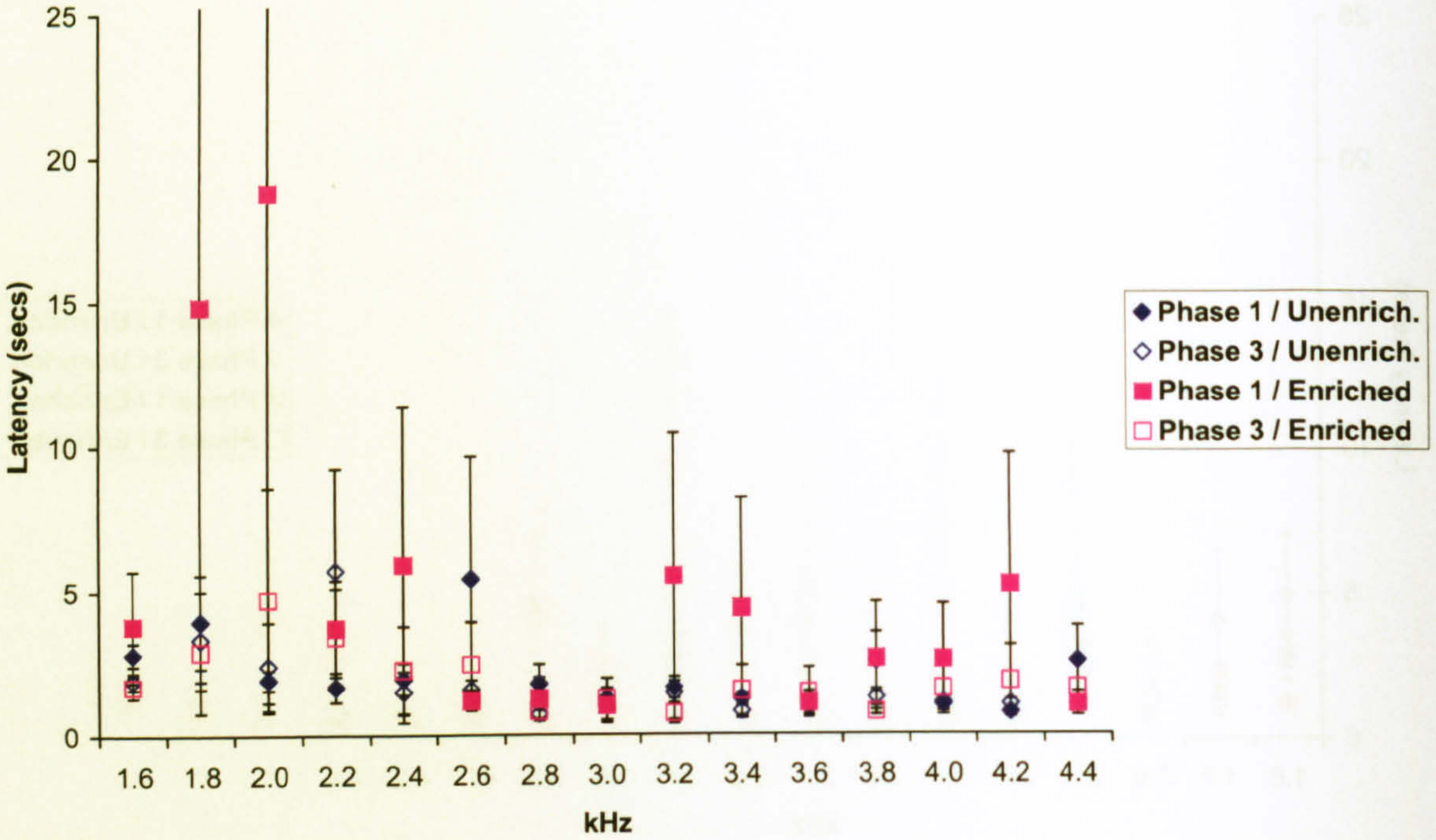

Figure 3.23 '1-pellet' lever latency: *2kHz=2pell / Measurement Phase 1 & 3*. As Figure 3.22, but plotting the first and last *measurement phase*.

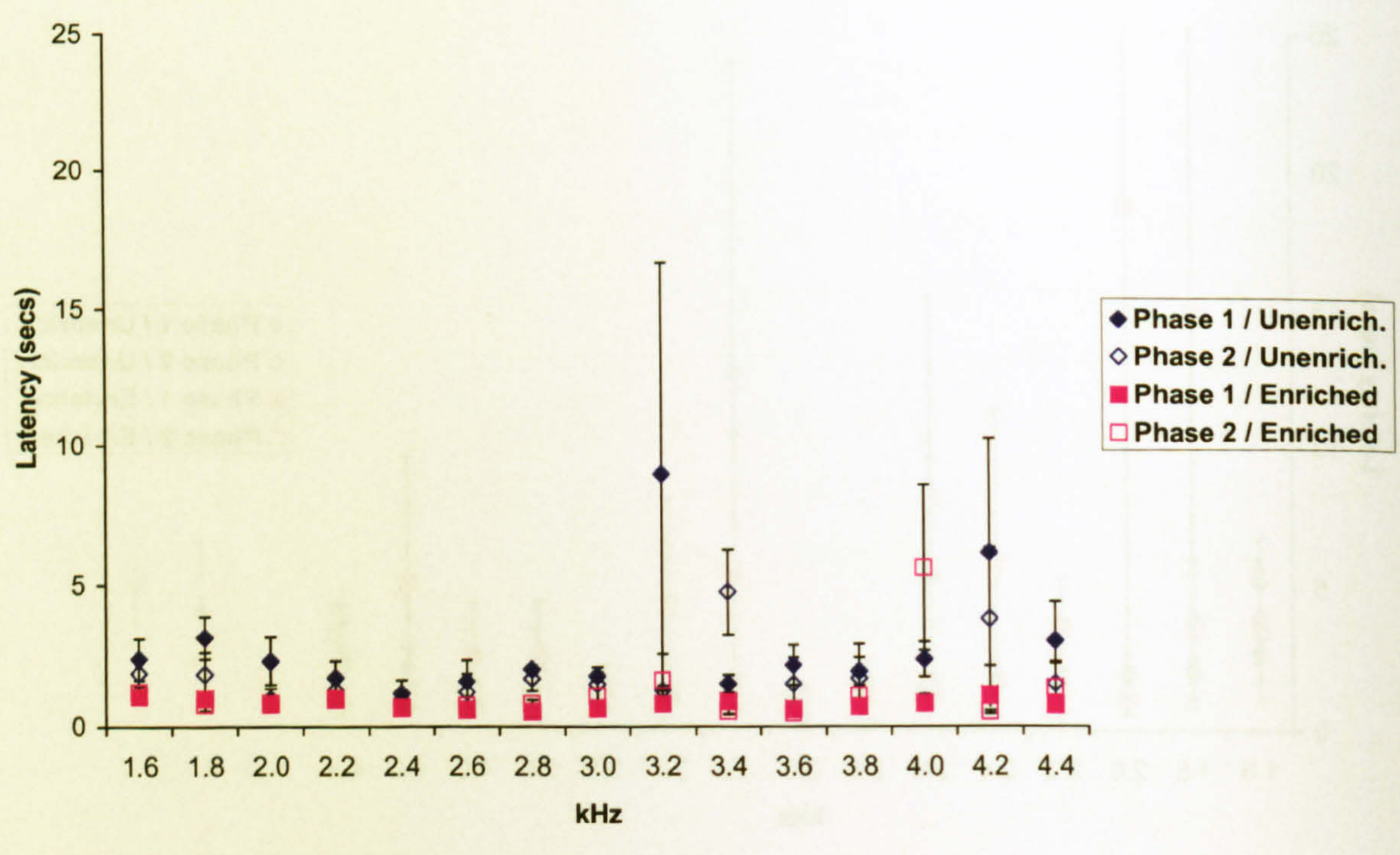

Figure 3.24 '1-pellet' lever latency: *4kHz=2pell / Measurement Phase 1 & 2*. As Figure 3.22, but for the *4kHz=2pell contingency* group.

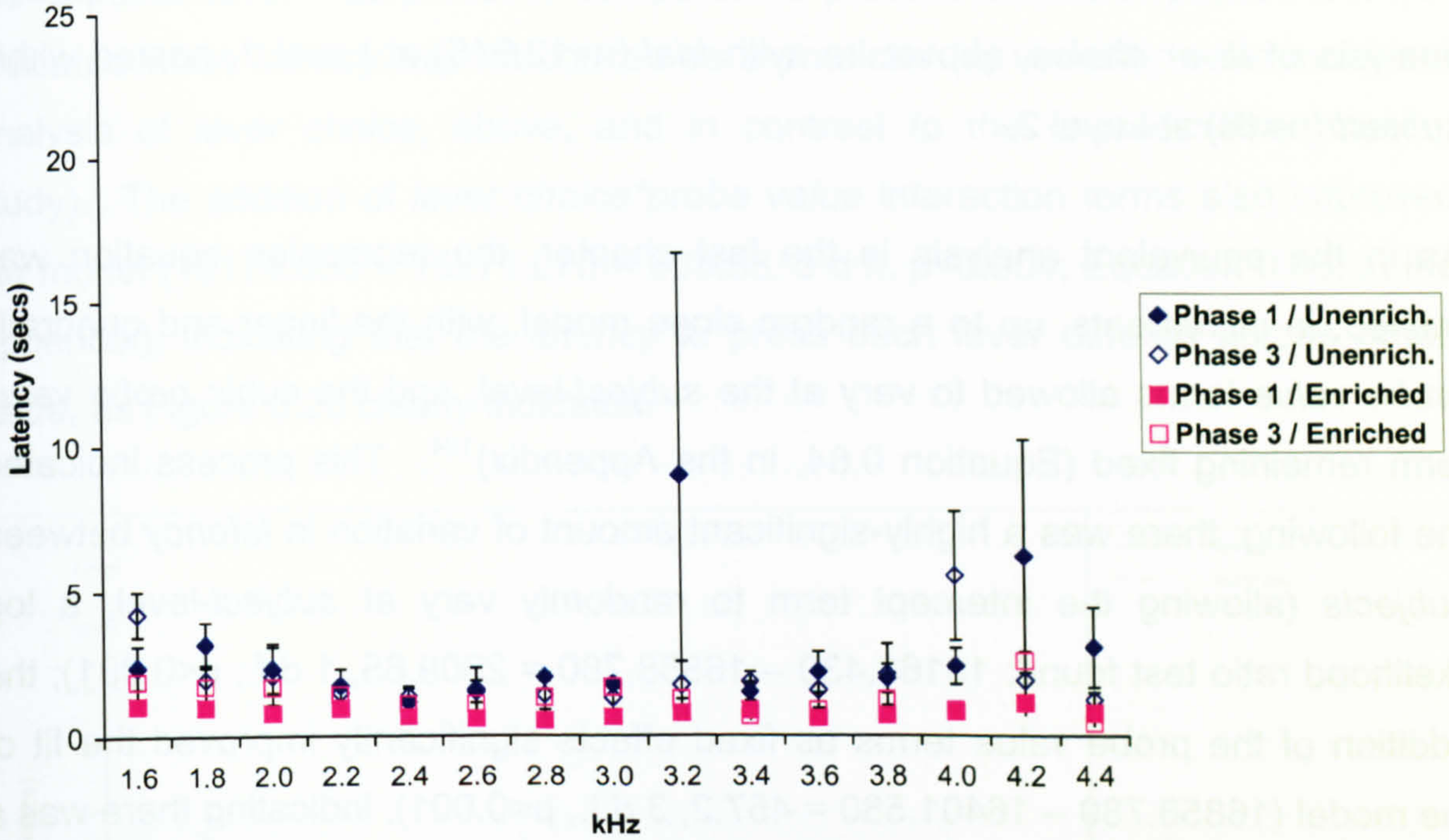

Figure 3.25 '1-pellet' lever latency: *4kHz=2pell / Measurement Phase 1 & 3.* As Figure 3.24, but plotting the first and last *measurement phases*.

Multi-level general linear model in MLwiN

As in the equivalent analysis in the last chapter, since histograms of the raw *latency* data revealed a marked positive skew, some form of transformation was likely to be necessary to meet the assumptions of a general linear model. The residuals from a variety of exploratory models were therefore inspected, and a negative reciprocal root transformation was once more chosen as the most appropriate transformation. When employing this transformation in the last chapter, a few outlying datapoints (n=7) of very short latency (≤0.12 seconds) were omitted; here, since there were a slightly larger number of such datapoints (n=24 at ≤0.12 seconds), rather than omit them, we rounded the *latency* data into slightly larger 'bins' (i.e. measurement units)[123] prior to transformation. See later for a further examination of whether the data met the assumptions of the model.

---

[123] The user can specify whether the Graphic State software, used to record the rats' responses in the operant chamber, samples the operant environment (and thus detects any changes, such a lever press) every 0.02, 0.05, or 0.1 seconds. We

The hierarchy of the dataset was defined in the same way as in the multilevel analysis of *lever choice*, above: i.e. with *trial* (n=12,915) at Level 1, nested within *subject* (n=16) at Level 2.

As in the equivalent analysis in the last chapter, the regression equation was refined, in increments, up to a random slope model, with the linear and quadratic *probe value* terms allowed to vary at the *subject*-level, and the cubic *probe value* term remaining fixed (Equation 0.64, in the Appendix)[124]. This process indicated the following: there was a highly-significant amount of variation in *latency* between *subjects* (allowing the intercept term to randomly vary at *subject*-level, a log likelihood ratio test found: 19168.430 – 16858.780 = 2309.65, 1 d.f., $p<0.001$); the addition of the probe value terms as fixed effects significantly improved the fit of the model (16858.780 – 16401.580 = 457.2, 3 d.f., $p<0.001$), indicating there was a shorter latency to record a lever response as the linear *probe value* scale increased in value (i.e. as it approached the end of the scale where the '2-pellet' reference stimulus was located); and there was a significant amount of *subject*-level variance in *latency* across *probe value* (allowing the linear and quadratic *probe value* terms to randomly vary across *subject:* 16401.580 – 16211.390 = 190.19, 5 d.f., $p<0.001$). The variance partition coefficients (VPC) we calculated from these relatively simple models indicated around 17 to 17.5%[125] of the overall variance could be attributed between-*subject* variation.

The addition of *Lever choice* (i.e. whether the lever associated with 2 pellets of food was pressed or not), as a main effect,[126] also improved the model's fit (16211.390 – 16174.850 = 36.54, 1 d.f., $p<0.001$; Equation 0.65, in the Appendix): the negative valence of the coefficient indicated there was a shorter *latency* when

---

originally chose a sampling interval (i.e. 'bin') of 0.02 seconds, but prior to the current analysis (and transformation) rounded this data to one decimal point (i.e. converted the bins to 0.1 seconds), with a minimum *latency* of 0.2 seconds. Whilst this resulted in a slight loss of information, this simple 'smoothing' procedure (e.g. Quinn & Keough, 2002) enabled us to model the full dataset whilst reducing the risk that very low latency outliers would disproportionately influence the model.

[124] As mentioned previously in the thesis, while describing the model we occasionally refer to equations copied into the Appendix: it's not necessary to refer to all (or any) of these, they are presented for purposes of optional reference, should the reader wish to check other aspects of the model not reported here.

[125] Calculated first without, and then with, the *probe value* terms, respectively. These values are similar to those calculated in the equivalent analysis in the last chapter (we found 17.4%, and 18%, respectively, in that analysis).

[126] Presses on the '1-pellet' lever were assigned the reference category, with a value of *'0'*, whilst presses on the '2-pellet' lever were assigned a value of *'1'*.

the '2-pellet' lever was pressed, compared to presses on the '1-pellet' lever (as indicated when *latency* was introduced as a predictor (x) variable in the multilevel analysis of *lever choice*, above, and in contrast to the unpredictable housing study). The addition of *lever choice*probe value* interaction terms also improved the model (16174.850 – 15271.270 = 903.58, 3 d.f., p<0.001; Equation 0.66, in the Appendix), indicating that the *latency* to press each lever differed across *probe value*, as Figure 3.26 clearly indicates.

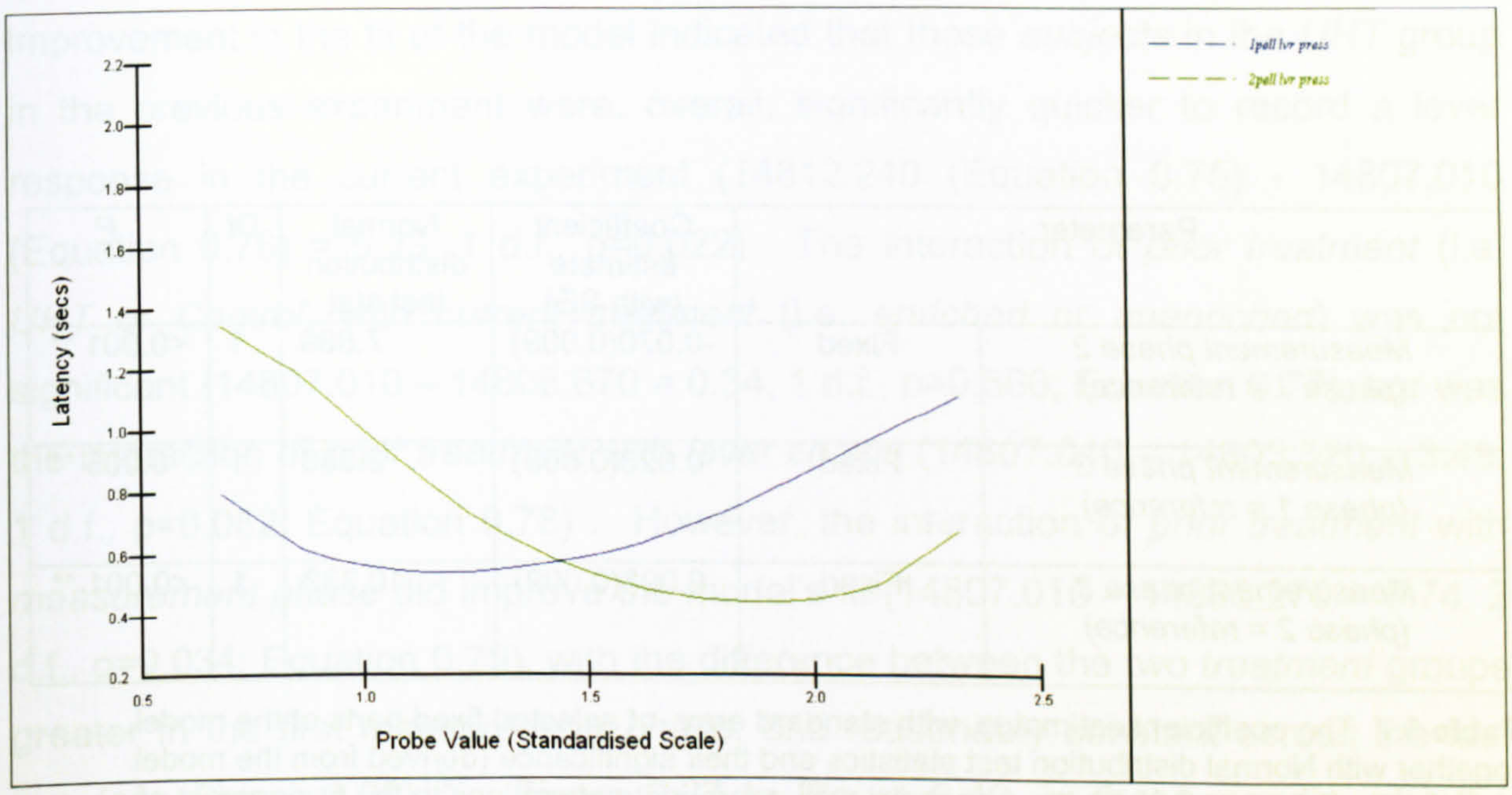

Figure 3.26 The predicted probability of pressing each lever, across *probe value* (these predictions were generated from the model specified in Equation 0.66, in the Appendix).

The addition of *Contingency* to the model revealed it did not have a significant main effect (15271.270 – 15269.700 = 1.57, 1 d.f., p=0.210; Equation 0.67), indicating there was no overall difference in *latency* across *contingency*. However, fitting *contingency* up to a three-way interaction with *lever choice* and *probe value*, did improve the fit of the model, and so these terms remained (15269.700 – 15035.500 = 234.2, 7 d.f., p<0.001; Equation 0.68).

The addition of *measurement phase*, as a main effect, made a very significant improvement to the fit of the model (15035.500 – 14926.610 = 108.89, 2 d.f., p<0.001; Equation 0.69, in the Appendix). As Table 3.7 indicates, Normal distribution tests found that the lever press responses in *Phase 2* were significantly

quicker than those in each of the other two *measurement phases*, whilst those in *Phase 1* were significantly quicker than those in *Phase 3* (i.e. the order of latency across *measurement phase,* with the quickest listed first, was: *Phase 2* < *Phase 1* < *Phase 3*). Various interactions of *measurement phase* with *probe value, lever choice* and *contingency* were explored: none improved the model's fit, bar the interaction of *measurement phase* with *contingency* (14926.610 – 14912.250 = 14.36, 2 d.f., p<0.001; Equation 0.70), and so this alone remained, along with the main effect.

Parameter		Coefficient estimate (with SE)	Normal distribution test stat	Df	*P*
Measurement phase 2 (phase 1 = reference)	Fixed	-0.070(0.009)	7.889	1	<0.001 **
Measurement phase 3 (phase 1 = reference)	Fixed	0.023(0.009)	2.556	1	0.005 **
Measurement phase 3 (phase 2 = reference)	Fixed	0.093(0.009)	10.333	1	<0.001 **

Table 3.7 The coefficient estimates, with standard error, of selected fixed parts of the model, together with Normal distribution test statistics and their significance (derived from the model specified in Equation 0.69, in the Appendix, with appropriate alterations to the assignment of reference category).

The addition of *treatment* (i.e. *enriched* or *unenriched*), as a main effect,[127] did not improve the model (14912.250 - 14910.990 = 1.26, 1 d.f., p=0.262; Equation 0.71), however the two-way interaction between *treatment* and *lever choice* did (14910.990 – 14873.540 = 37.45, 1 d.f., p<0.001; Equation 0.72). An examination of the coefficients revealed that the *enriched treatment* group were faster to press the '1-pellet' lever, and slower to press the '2-pellet' lever, than the *unenriched treatment* group. The interaction of *treatment* with *measurement phase* also substantially improved the model's fit to the data (14910.990 – 14851.200 = 59.79,

[127] The *Unenriched treatment* group were assigned the reference category, with a value of *'0'*, whilst the *Enriched treatment* group were assigned a value of *'1'*.

2 d.f., p<0.001; Equation 0.73). Figure 3.27[128] plots the predicted effect of the *enriched treatment* on *latency* across *measurement phase,* contrasted against the *unenriched treatment* group, which are held constant at zero. Both this chart, and Table 3.8, indicate that whilst the *enriched treatment* group are notably faster to record a response in the first two *measurement phases,* this difference is diminished in the final *measurement phase.*

Finally, *prior treatment* was added to the equation, firstly as a main effect.[129] The improvement in the fit of the model indicated that those *subjects* in the *UHT* group in the previous experiment were, overall, significantly quicker to record a lever response in the current experiment (14812.240 (Equation 0.75) - 14807.010 (Equation 0.76) = 5.23, 1 d.f., p=0.022). The interaction of *prior treatment* (i.e. *UHT* or *Control*) with *current treatment* (i.e. *enriched* or *unenriched*) was not significant (14807.010 – 14806.670 = 0.34, 1 d.f., p=0.560; Equation 0.77), nor was the interaction of *prior treatment* with *lever choice* (14807.010 – 14803.520 = 3.49, 1 d.f., p=0.062; Equation 0.78) . However, the interaction of *prior treatment* with *measurement phase* did improve the model's fit (14807.010 – 14800.270 = 6.74, 2 d.f., p=0.034; Equation 0.79), with the difference between the two *treatment* groups greater in the first *measurement phase*, and reasonably constant across the last two (see Figure 3.28)[130]. Figure 0.37 to Figure 0.40, in the Appendix, plot the residuals from this model, revealing the model's assumptions have been reasonably well upheld.

---

[128] Again, see Figure 0.35, in the Appendix, for an alternative way of plotting such predictions, as derived from a slightly different equation.

[129] The *Control (predictable housing)* group were assigned the reference category, with a value of *'0'*, whilst the *unpredictable housing treatment (UHT)* group were assigned a value of *'1'*.

[130] Or Figure 0.36, in the Appendix, for different method of plotting such data, using predictions derived from a slightly different equation.

Parameter		Coefficient estimate (with SE)	Normal distribution test stat	Df	*P*
*Treat.*Measurement phase 2 (phase 1 = reference)*	Fixed	-0.034(0.018)	1.889	1	0.029 *
*Treat.*Measurement phase 3 (phase 1 = reference)*	Fixed	0.103(0.018)	5.722	1	<0.001 **
*Treat.*Measurement phase 3 (phase 2 = reference)*	Fixed	0.137(0.018)	7.611	1	<0.001 **

Table 3.8 The coefficient estimates, with standard error, of selected fixed parts of the model, together with Normal distribution test statistics and their significance (derived from the model specified in Equation 0.73, in the Appendix, with appropriate alterations to the assignment of reference category).

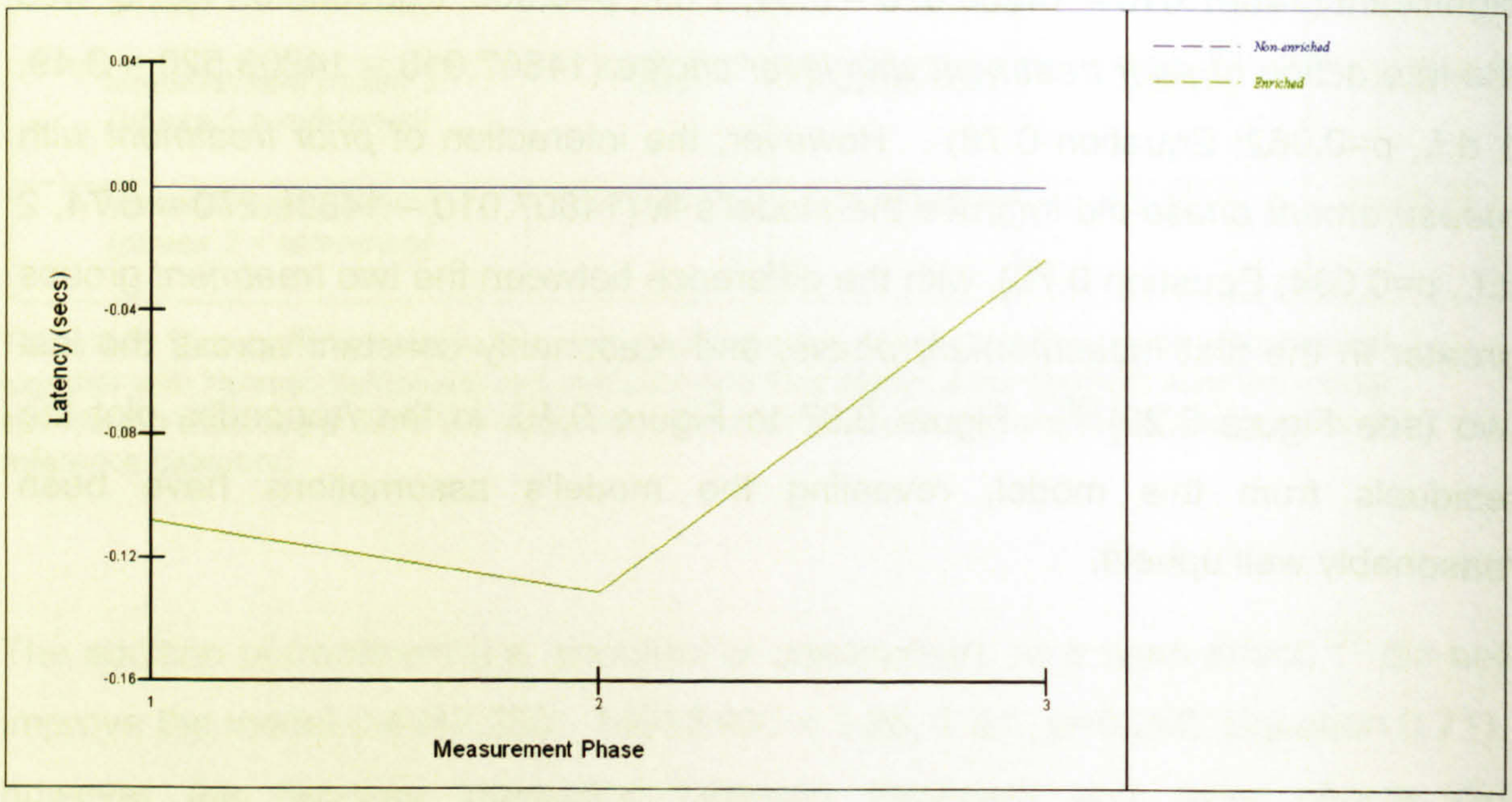

Figure 3.27 The predicted effect of the *enriched treatment* on the *latency* to press either lever, across different levels of *measurement phase* (from a prediction equation derived from the model specified in Equation 0.73). Here, the *enriched* group is compared against the reference category of the *unenriched* group, which has a *latency* held constant at zero seconds (see Figure 0.34, in the Appendix, for an alternative way of plotting such data).

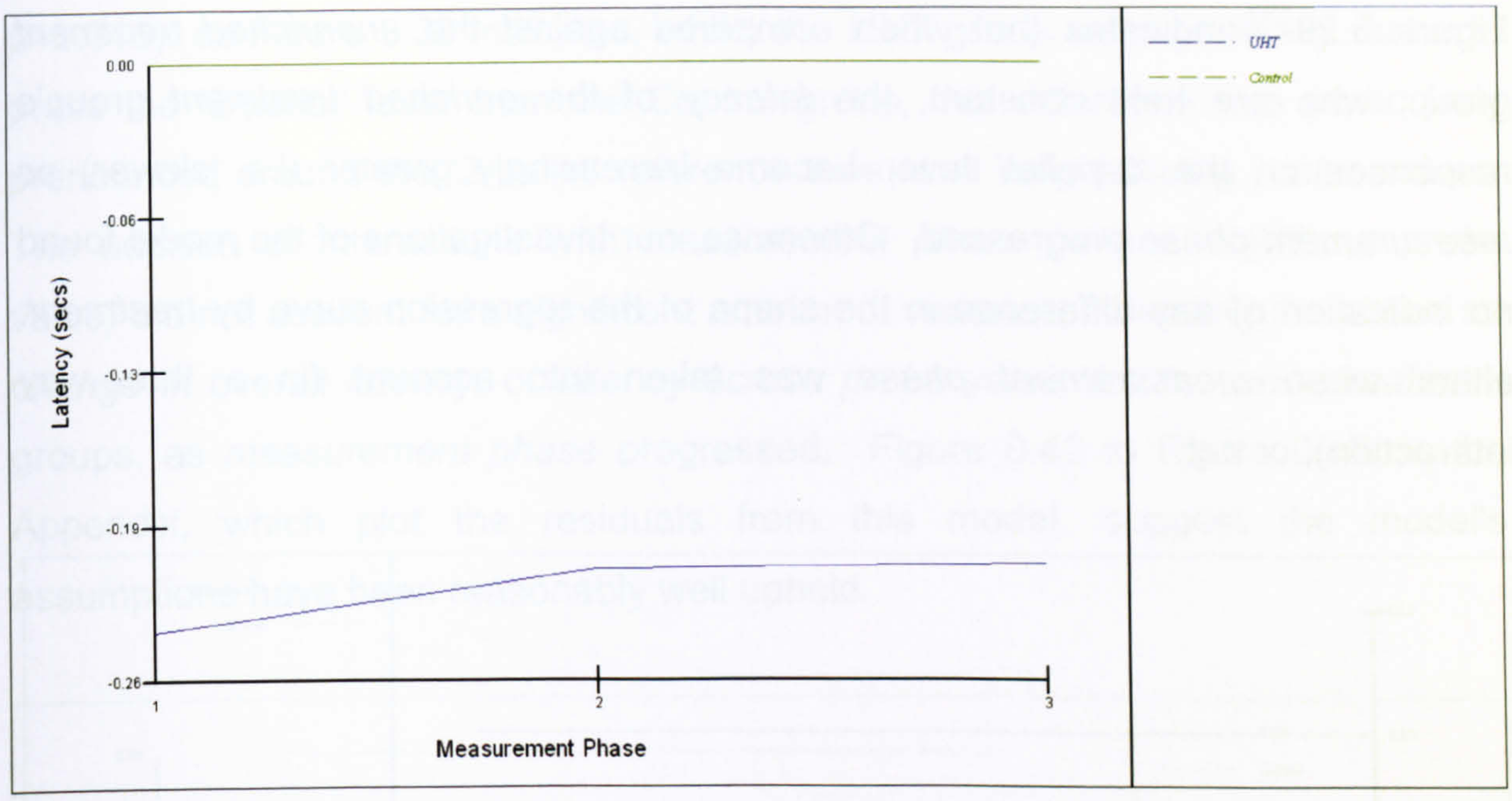

Figure 3.28 The predicted effect of the *UHT (prior) treatment* on the *latency* to press either lever, across different levels of *measurement phase* (from a prediction equation derived from the model specified in Equation 0.79). Here, the *UHT* group is compared against the reference category of the *Control (predictable housing)* group, which has a *latency* held constant at zero seconds (see Figure 0.36, in the Appendix, for an alternative way of plotting such data).

Finally, as in the equivalent analysis in the last chapter, we briefly review two simpler, subsidiary models, each of which only models the *latency* data from one lever (i.e. the '2-pellet' or '1-pellet' lever). For the categorical predictor (x) variables, the choice of reference category remained the same as for the *latency* analysis of all the lever presses, above.

The analysis modelling only those presses made on the '2-pellet' lever (n=6,436)[131] found no significant main effect of *treatment* (i.e. *enriched* or *unenriched*) on the *latency* to record a lever press (6377.846 (Equation 0.81) – 6375.550 (Equation 0.82) = 2.296, 1 d.f., p=0.130), although the addition of the interaction between *treatment* and *measurement phase* did substantially improve the model's fit (6375.550 (Equation 0.82) – 6341.172 (Equation 0.83) = 34.378, 2 d.f., p<0.001);

[131] Due to convergence problems encountered during model-building, the model was specified slightly different in this analysis, with the quadratic term for *probe value* being fixed (i.e. not allowed to vary randomly at the *subject*-level). When we compared this model to those in which the quadratic term was allowed to randomly vary, and which successfully converged, the conclusions presented here, in the main text, remained the same.

Figure 3.29[132] indicates that when compared against the *unenriched treatment* group, who are held constant, the *latency* of the *enriched treatment* group's responses on the '2-pellet' lever became increasingly greater (i.e. slower) as *measurement phase* progressed. Otherwise, our investigations of the model found no indication of any difference in the shape of the regression curve by *treatment*: either when *measurement phase* was taken into account (in a three-way interaction), or not.

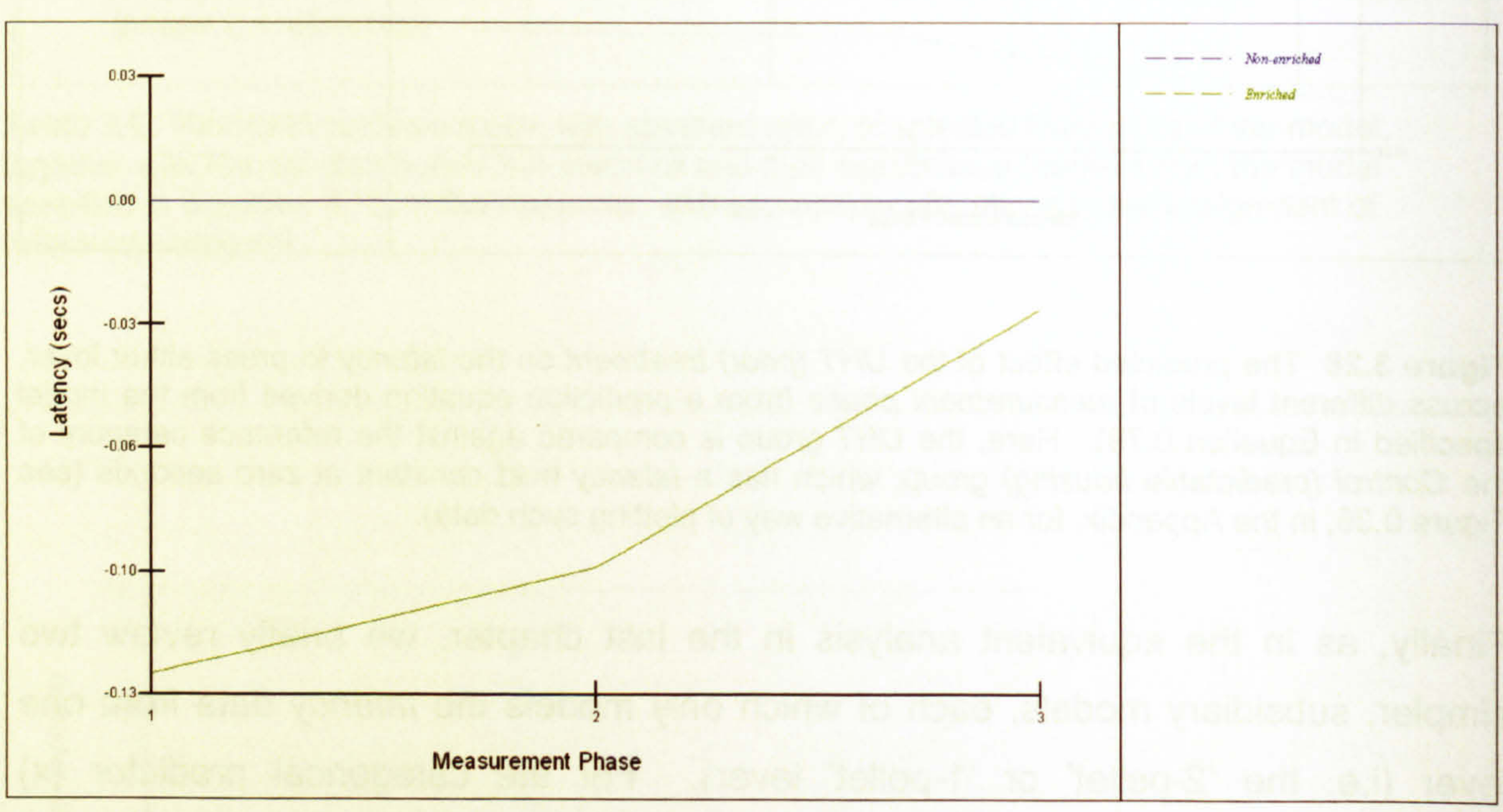

Figure 3.29 The predicted effect of the *enriched treatment* on the *latency* to press the '2-pellet' lever, across different levels of *measurement phase* (from a prediction equation derived from the model specified in Equation 0.83). Here, the *enriched treatment* group is compared against the reference category of the *unenriched treatment* group, which has a *latency* held constant at zero seconds (see Figure 0.41, in the Appendix, for an alternative way of plotting such data).

The addition of *prior treatment,* as a main effect, indicated that the *unpredictable housing treatment (UHT)* group were significantly faster to press the '2-pellet' lever (6341.172 (Equation 0.83) – 6332.986 (Equation 0.85) = 8.186, 1 d.f., p=0.004). The interaction between *prior treatment* and *probe value* also improved the model's fit (6332.986 (Equation 0.85) – 6321.554 (Equation 0.86) = 11.432, 3 d.f.,

[132] See Figure 0.41, in the Appendix, for different method of plotting such data, using predictions derived from a slightly different equation.

p=0.010): as Figure 3.30 indicates, whilst the *UHT* group were generally faster to press the '2-pellet' lever than the *Control* group, this difference was particularly pronounced around the '2-pellet' reference tone, and neighbouring *probe values.* The addition of interactions with *measurement phase* (with and without *probe value)* did not account for a significant amount of variance, indicating there was no change in overall *latency,* or *latency* across *probe value,* for the *prior treatment* groups, as *measurement phase* progressed. Figure 0.42 to Figure 0.45, in the Appendix, which plot the residuals from this model, suggest the model's assumptions have been reasonably well upheld.

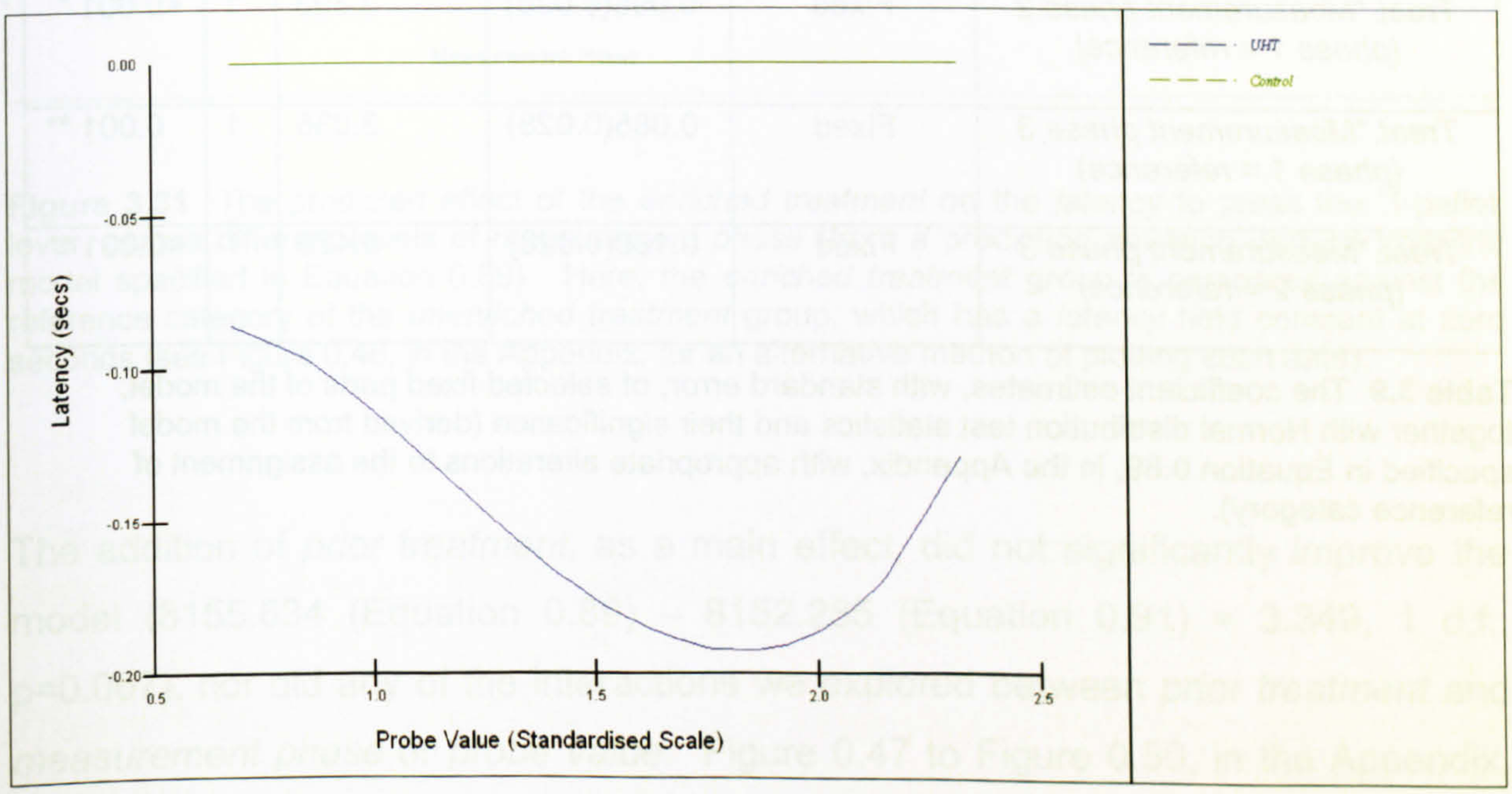

Figure 3.30 The predicted effect of the *UHT (prior) treatment* on the *latency* to press the '2-pellet' lever, across *probe value* (a standardised scale, with *'1'* corresponding to the reference stimulus associated with 1 pellet of food, and *'2'* corresponding to the reference stimulus associated with 2 pellets of food. Prediction equation derived from the model specified in Equation 0.86). Here, the *UHT* group is compared against the reference category of the *Control (predictable-housing)* group, which has a *latency* held constant at zero seconds.

The analysis modelling only those presses made on the '1-pellet' lever (n=6,479) found a significant main effect of *treatment,* with the *enriched treatment* group significantly quicker to press the '1-pellet' lever (8207.634 (Equation 0.87) - 8199.477 (Equation 0.88)= 8.157, 1 d.f., p=0.004). The two-way interaction of *treatment* with *measurement phase* made a highly-significant contribution to the model (8199.477 (Equation 0.88) - 8155.634 (Equation 0.89) = 43.843, 2 d.f.,

$p<0.001$): as Table 3.9 and Figure 3.31[133] indicate, whilst the *enriched treatment* group were consistently faster to press the '1-pellet' lever, this difference increased from the first to second *measurement phase,* and decreased in the final *measurement phase.* The interactions of *treatment* with *probe value,* either with or without *measurement phase,* did not contribute significantly to the model, indicating no difference in the *latency* to press the '1-pellet' lever between the two *treatment* groups depending on which *probe value* has been presented.

Parameter		Coefficient estimate (with SE)	Normal distribution test stat	Df	*P*
*Treat.*Measurement phase 2 (phase 1 = reference)*	Fixed	-0.095(0.028)	3.393	1	<0.001 *
*Treat.*Measurement phase 3 (phase 1 = reference)*	Fixed	0.085(0.028)	3.036	1	0.001 **
*Treat.*Measurement phase 3 (phase 2 = reference)*	Fixed	0.180(0.028)	6.429	1	<0.001 **

Table 3.9 The coefficient estimates, with standard error, of selected fixed parts of the model, together with Normal distribution test statistics and their significance (derived from the model specified in Equation 0.89, in the Appendix, with appropriate alterations to the assignment of reference category).

[133] See Figure 0.46, in the Appendix, for different method of plotting such data, using predictions derived from a slightly different equation.

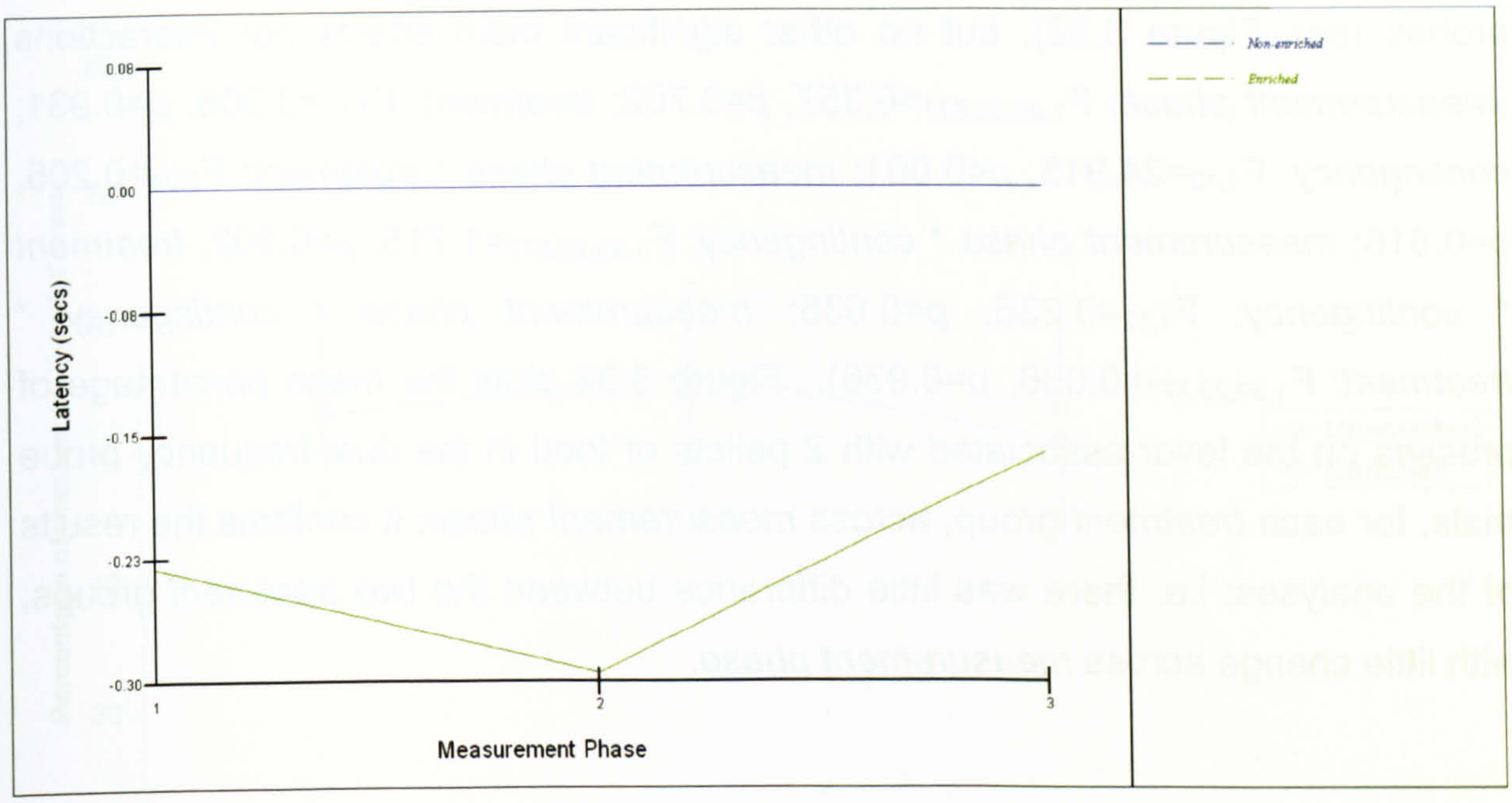

Figure 3.31 The predicted effect of the *enriched treatment* on the *latency* to press the '1-pellet' lever, across different levels of *measurement phase* (from a prediction equation derived from the model specified in Equation 0.89). Here, the *enriched treatment* group is compared against the reference category of the *unenriched treatment* group, which has a *latency* held constant at zero seconds (see Figure 0.46, in the Appendix, for an alternative method of plotting such data).

The addition of *prior treatment,* as a main effect, did not significantly improve the model (8155.634 (Equation 0.89) – 8152.285 (Equation 0.91) = 3.349, 1 d.f., p=0.067), nor did any of the interactions we explored between *prior treatment* and *measurement phase* or *probe value.* Figure 0.47 to Figure 0.50, in the Appendix, plot the residuals from the model specified in Equation 0.89, and suggest the model's assumptions have been reasonably well upheld.

Lever choice in dual-frequency probe trials

The mean percentage of presses on the lever associated with 2 pellets of food following presentation of the probe stimuli in each dual-frequency probe session was taken for each *subject*, and analysed in a repeated-measures GLM, with *measurement phase* as a within-subjects factor, and *contingency* and *treatment* as between-subjects factors. There was a highly-significant main effect of *contingency*, indicating that the *4kHz=2pell* group were more likely to press the lever associated with 2 pellets of food when presented with the dual-frequency

probes (see Figure 3.32), but no other significant main effects nor interactions (*measurement phase:* $F_{1.99,23.877}$=0.357, *p*=0.702; *treatment:* $F_{1,12}$=0.008, *p*=0.931; *contingency:* $F_{1,12}$=24.915, *p*<0.001; *measurement phase * treatment:* $F_{2,24}$=0.205, *p*=0.816; *measurement phase * contingency:* $F_{1.99,23.877}$=1.715, *p*=0.202; *treatment * contingency:* $F_{1,12}$=0.236, *p*=0.636; *measurement phase * contingency * treatment:* $F_{1.99,23.877}$=0.066, *p*=0.936). Figure 3.33 plots the mean percentage of presses on the lever associated with 2 pellets of food in the dual-frequency probe trials, for each *treatment* group, across *measurement phase*; it confirms the results of the analyses: i.e. there was little difference between the two *treatment* groups, with little change across *measurement phase*.

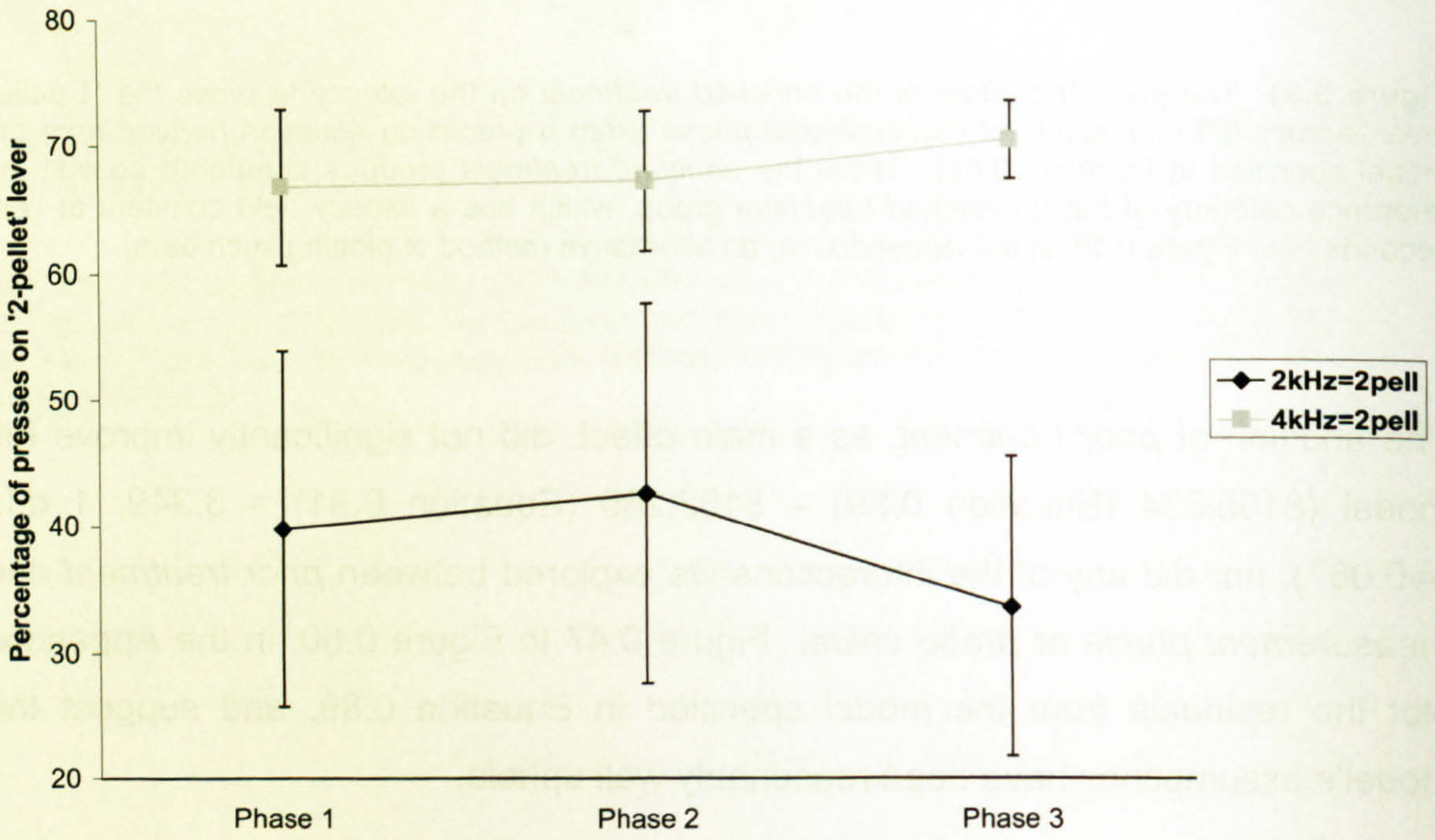

Figure 3.32 The mean percentage of presses on the '2-pellet' lever in the dual-frequency probe trials, across *measurement phase* for each *contingency* group (+/- 1SEM).

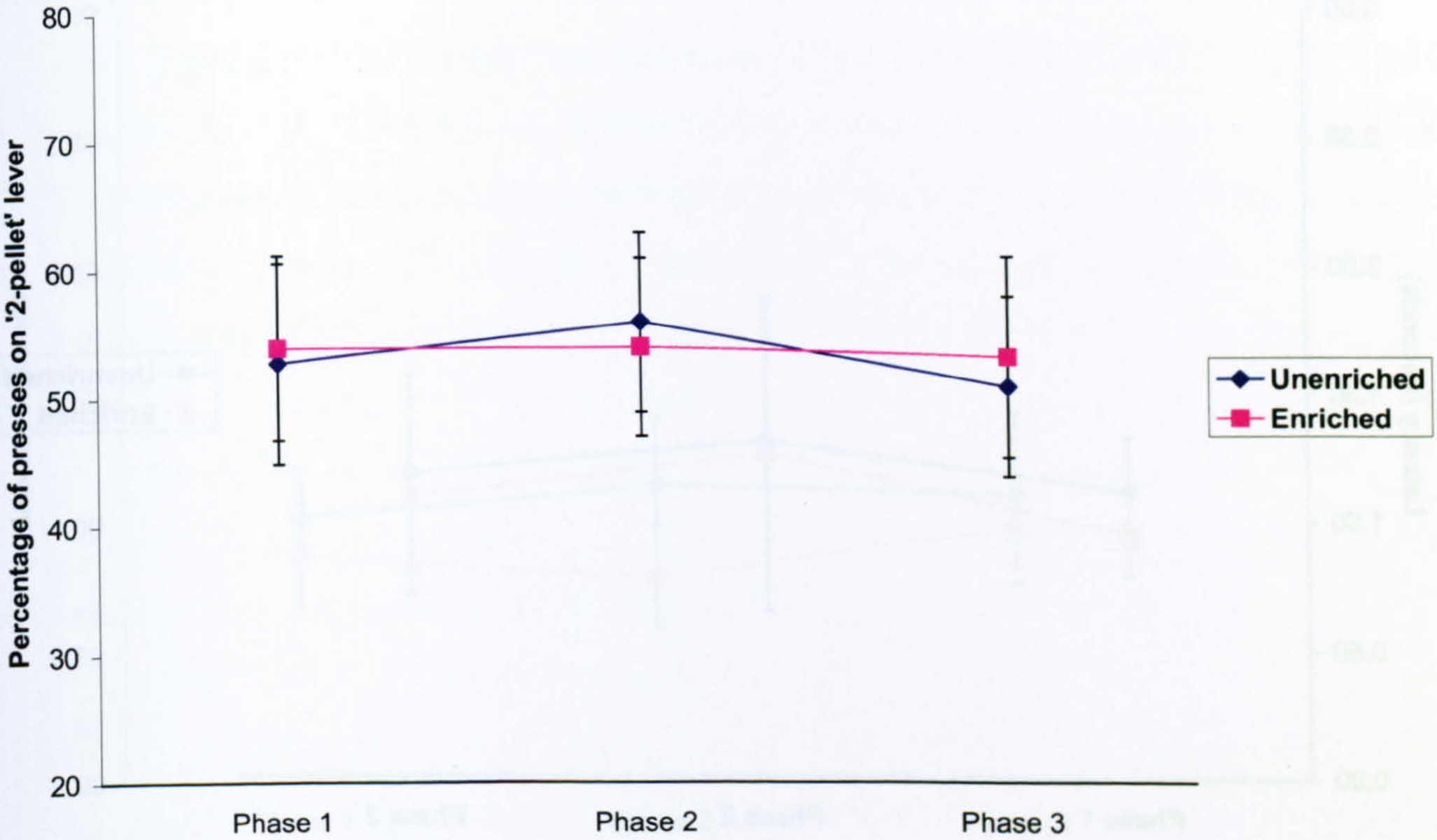

Figure 3.33 The mean percentages of presses made on the '2-pellet' lever in the dual-frequency probe trials, across *measurement phase* for each *treatment* group (+/- 1SEM).

Latency in dual-frequency probe trials

The mean *latency* to press either lever (i.e. regardless of that lever's identity) in the probe trials was taken for each *subject* for each dual-frequency test session, and analysed (following a negative inverse-transformation) in a repeated-measures GLM, with *measurement phase* as a within-subjects factor, and *contingency* and *treatment* as between-subjects factors; this found all main effects and interactions to be non-significant (*treatment:* $F_{1,12}$=0.097, p=0.761; *contingency:* $F_{1,12}$=2.052, p=0.178; *measurement phase:* $F_{1.735,20.816}$=0.733, p=0.474; *measurement phase * treatment:* $F_{1.735,20.816}$=0.666, p=0.504; *measurement phase * contingency:* $F_{1.735,20.816}$=0.020, p=0.970; *treatment * contingency:* $F_{1,12}$=0.048, p=0.831; *measurement phase * contingency * treatment:* $F_{1.735,20.816}$=0.166, p=0.819). Figure 3.34 plots the mean *latency* to press either lever in the dual-frequency probe trials for each *treatment* group, across *measurement phase*, confirming little difference between the groups, and little change across *measurement phase*.

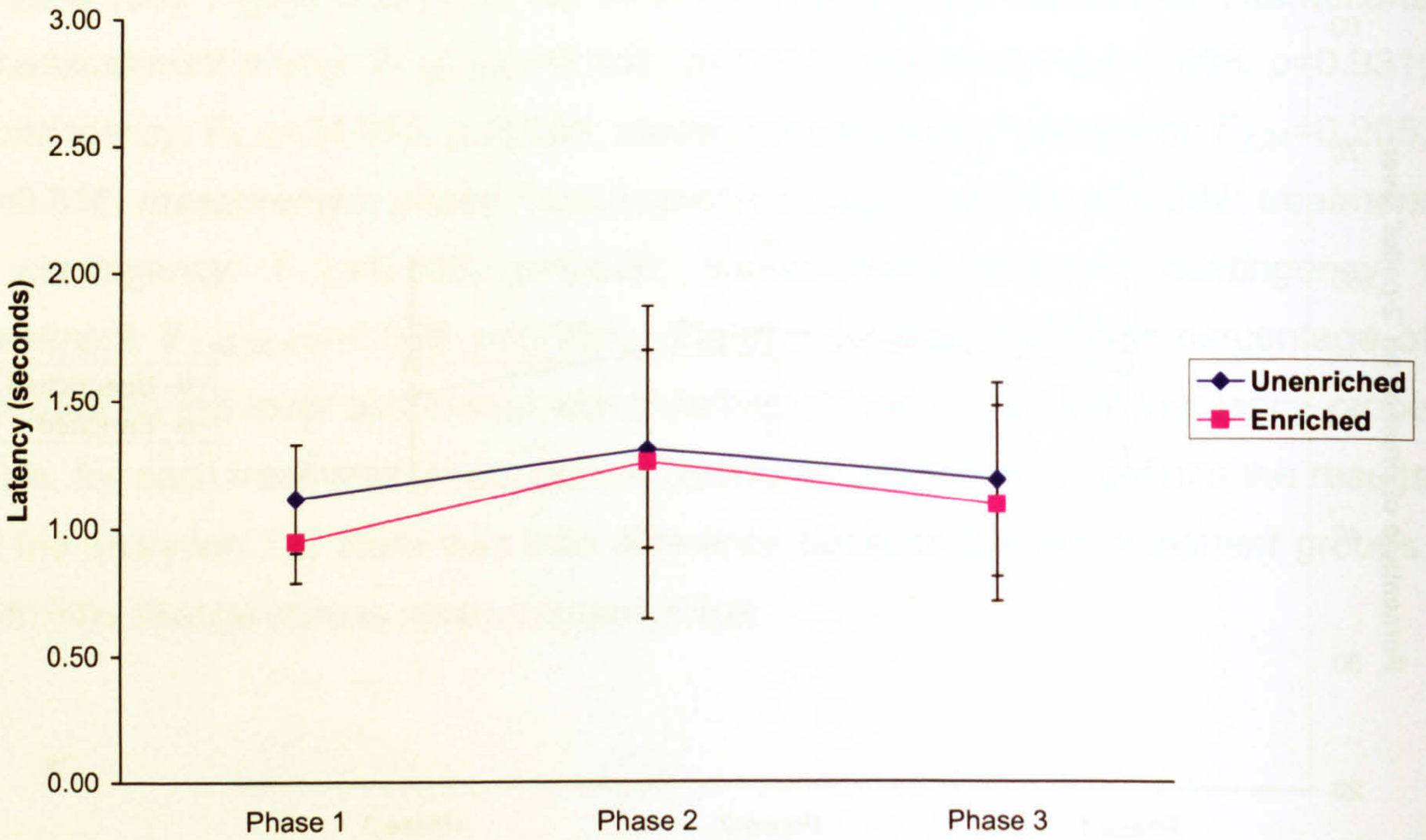

Figure 3.34 The mean latency to press either lever in the dual-frequency probe trials, across *measurement phase* for each *treatment* group (+/- 1SEM).

Additional analyses were conducted on the *latency* to press each *different* lever (i.e. as a function of whether it was associated with 1 or 2 pellets of food). The (log-transformed) *latency* to press the lever associated with 1 pellet of food in the dual-frequency probe trials was analysed using the same repeated-measures model; this found no significant main effect nor interactions (*measurement phase:* $F_{1.904,22.845}=0.136$, $p=0.864$; *treatment:* $F_{1,12}=0.484$, $p=0.500$; *contingency:* $F_{1,12}=0.598$, $p=0.454$; *measurement phase * treatment:* $F_{1.904,22.845}=0.572$, $p=0.564$; *measurement phase * contingency:* $F_{1.904,22.845}=0.294$, $p=0.738$; *treatment * contingency:* $F_{1,12}=0.473$, $p=0.505$; *measurement phase * contingency * treatment:* $F_{1.904,22.845}=1.074$, $p=0.355$). Figure 3.35 shows the change in mean *latency* across *measurement phase* for each of the *treatment* groups, confirming little difference between the groups, and little change across *measurement phase*.

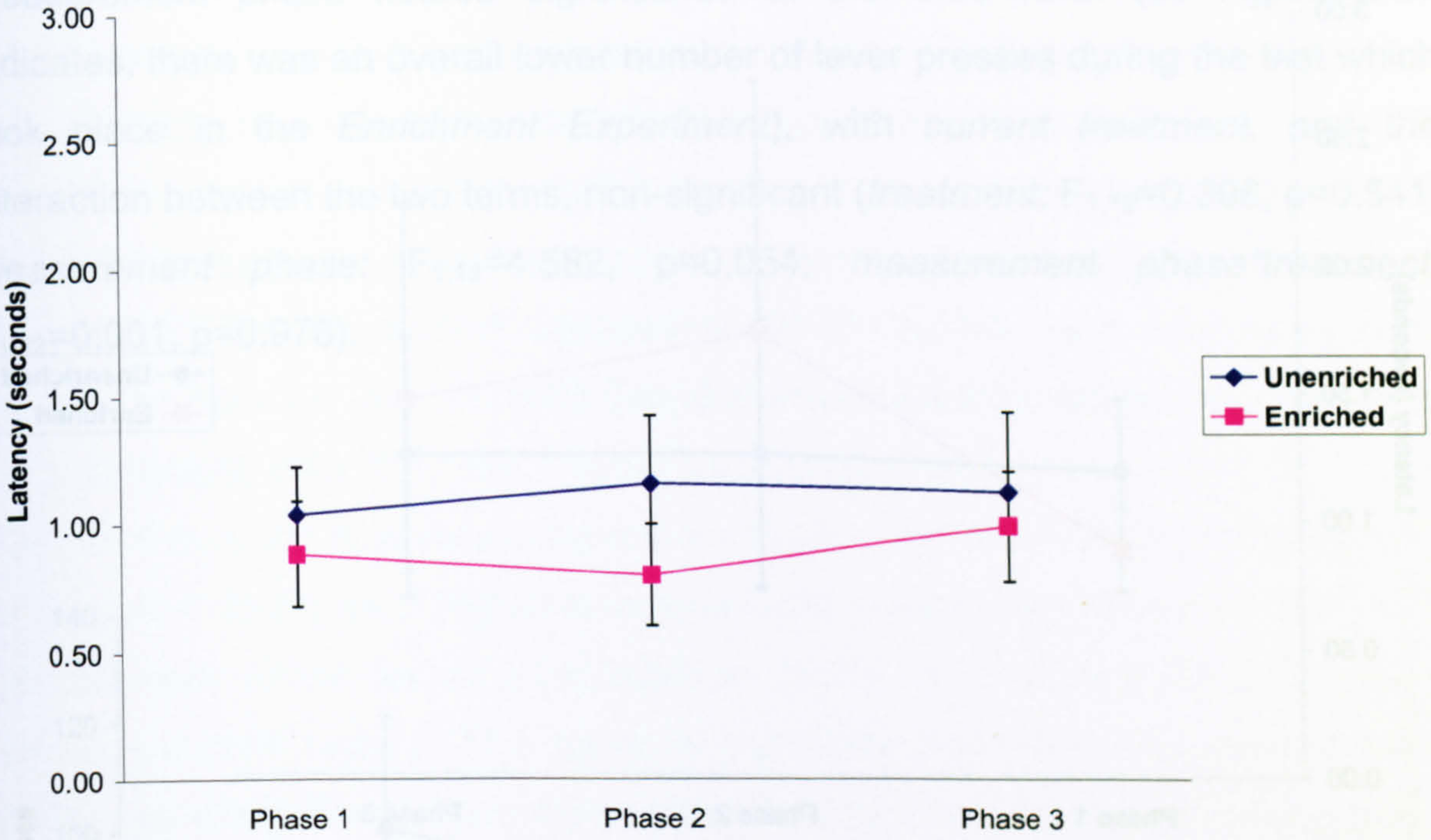

Figure 3.35 The mean latency to press the '1-pellet' lever in the dual-frequency probe trials, across *measurement phase* for each *treatment* group (+/- 1SEM).

The (negative inverse-transformed) *latency* to press the lever associated with 2 pellets of food in the dual-frequency probe trials was analysed in the same way[134]; this found no significant main effect nor interactions (*measurement phase:* $F_{1.960,23.516}=1.139$, $p=0.336$; *treatment:* $F_{1,12}=0.003$, $p=0.956$; *contingency:* $F_{1,12}=2.760$, $p=0.123$; *measurement phase * treatment:* $F_{1.960,23.516}=0.189$, $p=0.825$; *measurement phase * contingency:* $F_{1.960,23.516}=0.045$, $p=0.954$; *treatment * contingency:* $F_{1,12}=0.193$, $p=0.668$; *measurement phase * contingency * treatment:* $F_{1.960,23.516}=3.136$, $p=0.063$). Figure 3.36 shows the change in mean *latency* across *measurement phase* for each of the *treatment* groups.

[134] N.B. all assumptions of the GLM were met, bar the variances of the residuals from *Phase 1*, which failed a formal test of homogeneity (Levene's: p=0.002). While the log-transformed data faired better in the formal test, inspection of the residual plots suggested little improvement, and the negative-inverse transformation was chosen since it achieved better normality.

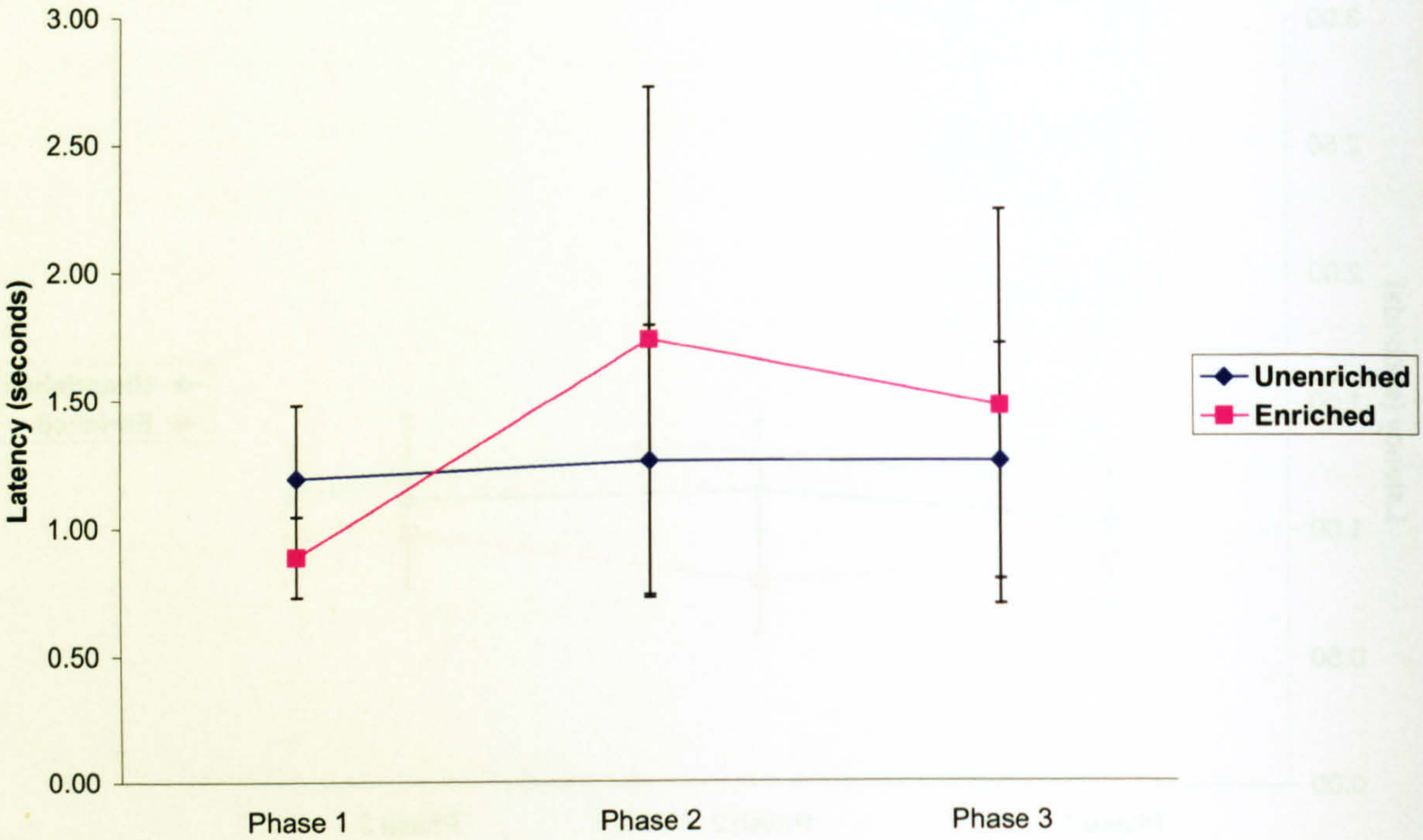

Figure 3.36 The mean latency to press the '2-pellet' lever in the dual-frequency probe trials, across *measurement phase* for each *treatment* group (+/- 1SEM).

Lever-based progressive ratio test with food reinforcement

Figure 3.37 plots the total number of times the lever was pressed during the progressive ratio test session by *current treatment* group. The data is plotted from both the current (environmental enrichment) experiment (the test took place 28 days after the change in enrichment), and, for reference, from the previous (unpredictable-housing) study. Of course, the rats were not in these *current treatment* groups during the unpredictable-housing study, and so are grouped by *future* treatment assignment, counterbalanced with respect to their *previous* treatment grouping (*UHT* or *Control*) in that experiment. Since the design was therefore orthogonal, the data were submitted to a repeated-measures GLM, with *measurement phase* (*UHT Experiment: Post-treatment* and *Enrichment Experiment: Post-treatment*) as the within-subjects factor, and *current treatment* (*enriched* and *unenriched*) as a between-subjects factor[135]. This found

---

[135] This analysis exluded all data from the subject (later assigned to the *unenriched treatment* group) that was excluded from the test during the unpredictable-housing study due to ill-health, and all data from the subject (*enriched treatment* group) euthanised shortly before the test in the environmental enrichment study.

measurement phase neared significance at the 0.05 level (as Figure 3.37 indicates, there was an overall lower number of lever presses during the test which took place in the *Enrichment Experiment*), with *current treatment,* and the interaction between the two terms, non-significant (*treatment:* $F_{1,12}$=0.396, p=0.541; *measurement phase:* $F_{1,12}$=4.582, p=0.054; *measurement phase*treatment:* $F_{1,12}$=0.001, p=0.976).

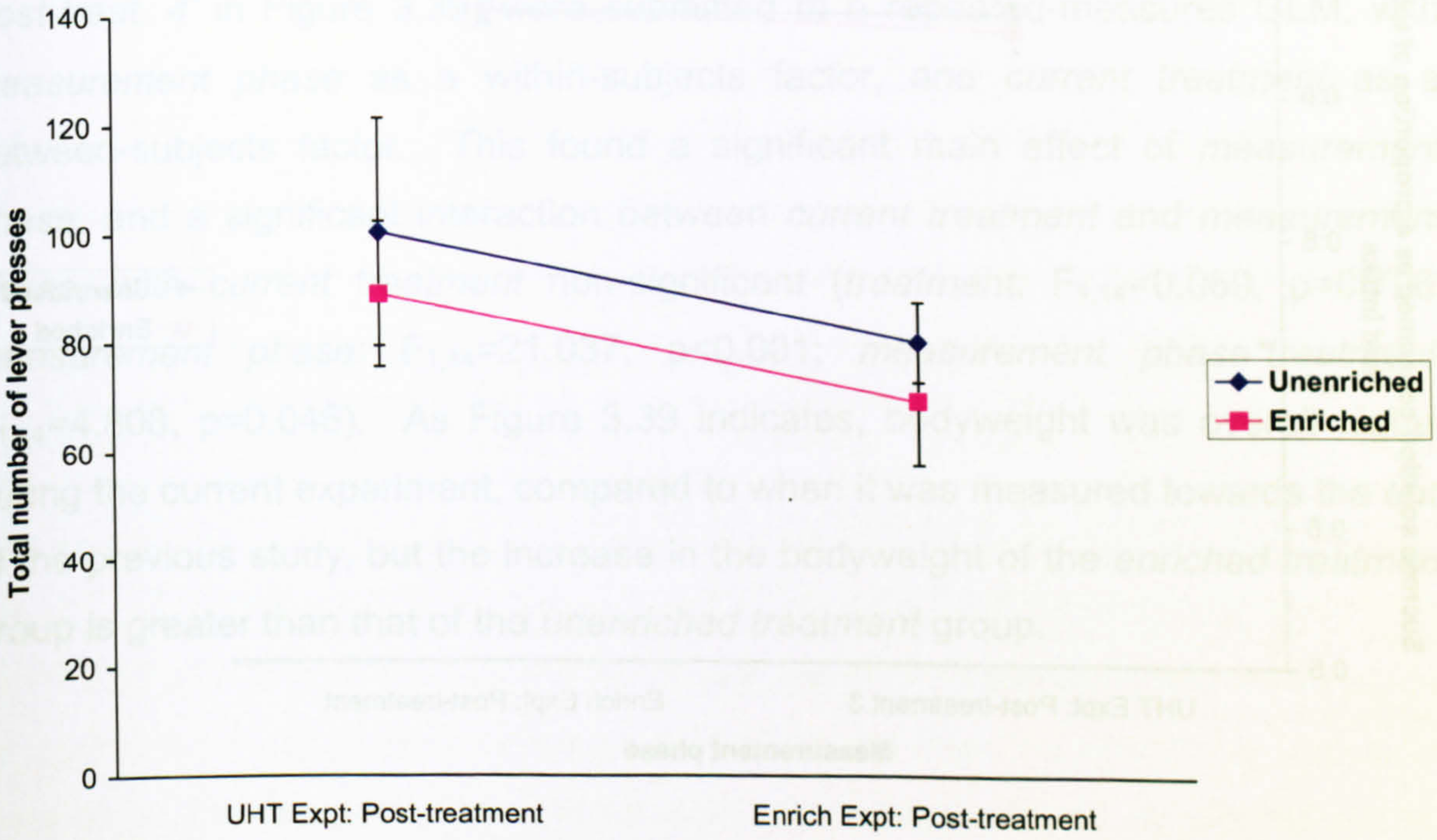

Figure 3.37 The mean number of times the lever was pressed in the progressive ratio test session by *treatment* group, across *measurement phase* (error bar = 1SEM). These summaries do not include any of the data from the two *subjects* (one in each of the *unenriched* and *enriched treatment* groups) that were excluded from one or other of the two test occasions due to non-experimental reasons (including their data does not substantially change the pattern of the plots).

Sucrose consumption: homecage-based test

Figure 3.38 plots the amount of sucrose solution consumed as a proportion of total fluid (i.e. water and sucrose solution) intake in each homecage, by *treatment* group. The results are plotted both from the sucrose consumption test conducted in the current (environmental enrichment) experiment, and also, for reference, from

the last occasion the test took place: in the preceding (unpredictable housing) experiment (see Figure 2.47). As the chart suggests, a repeated-measures GLM, with *treatment* as the between-subjects factor, and *measurement phase* as the within-subjects factor, found no significant differences in the amount of (arcsin-square-root-transformed) sucrose solution consumed as a proportion of total fluid intake (*treatment:* $F_{1,6}$=0.176, p=0.689; *measurement phase:* $F_{1,6}$=1.022, p=0.351; *measurement phase*treatment:* $F_{1,6}$=0.365, p=0.568).

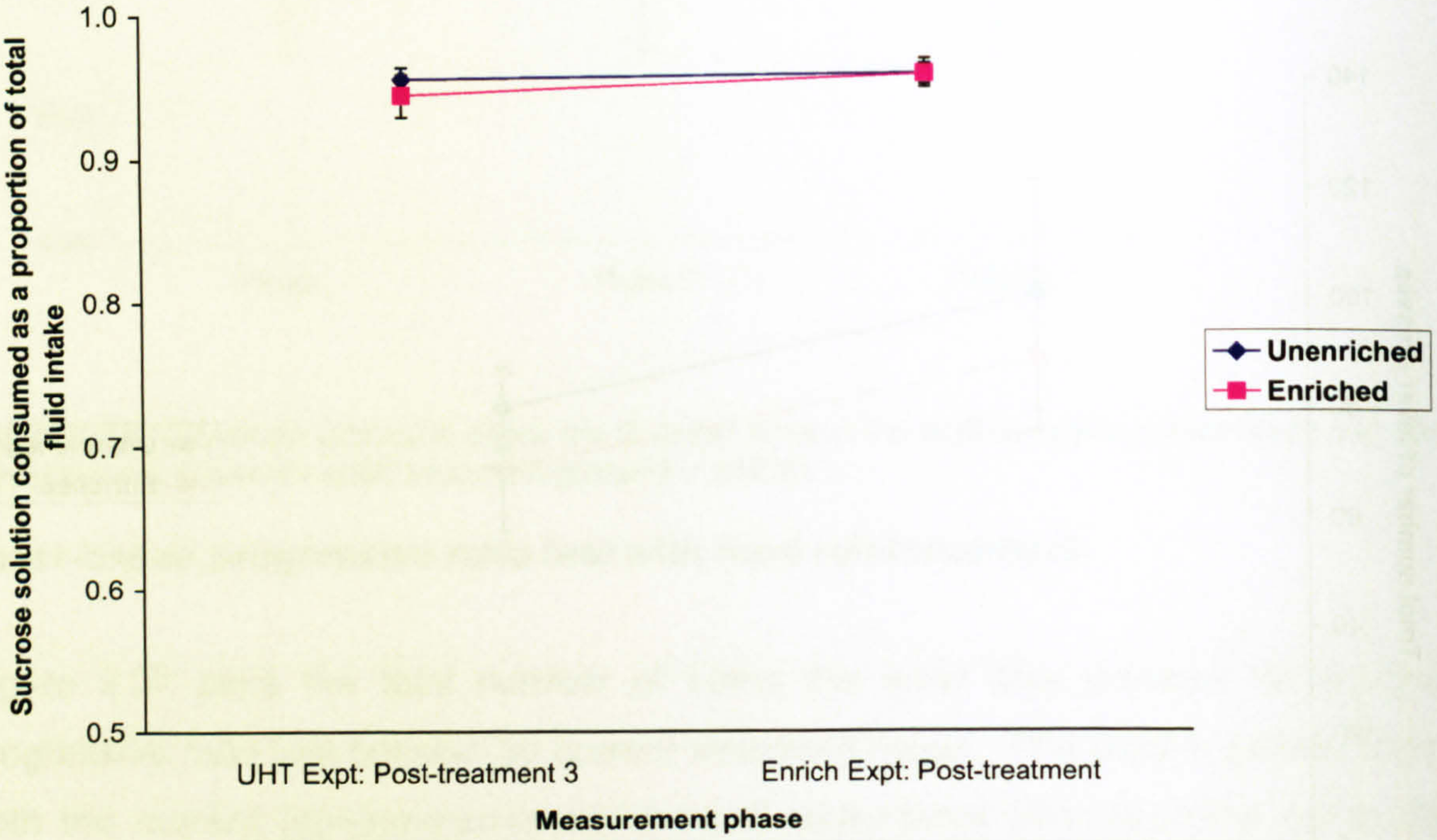

Figure 3.38 The mean amount of sucrose solution consumed as a proportion of total fluid intake (i.e. both water and sucrose solution) in a homecage-based sucrose preference test, by *treatment,* during different *measurement phases* (+/- 1SEM). On the x-axis, *'Enrich Expt: Post-treatment'* refers to the test which took place during the current (environmental enrichment) experiment, x days following a change in environmental enrichment. For reference, the results from the last occasion the rats received this test, towards the end of the previous (unpredictable-housing) experiment, are also plotted along the x-axis, as *'UHT Expt: Post-treatment 3'* (see Figure 2.47). All the data are summarised by *current treatment* group (*enriched* or *unenriched*): i.e. the data from *'UHT Expt: Post-treatment 3'* are grouped by *future* treatment assignment (in the unpredictable-housing experiment, the rats were in treatment groups specific to that study (*UHT* and *control*); these were counter-balanced across their later assignment to the *enriched* or *unenriched* groups).

Bodyweights

Figure 3.39 plots the rats' bodyweight, in grams, 27 days following the change in environmental enrichment, by *current treatment* group (i.e. *enriched* or *unenriched*). For reference, the bodyweight data, summarised by the same *current*

treatment grouping, are also plotted from the previous (unpredictable-housing) experiment. Of course, the rats were not in the *enriched* and *unenriched treatment* groups during the unpredictable-housing study: rather, they were in the groups specific to that experiment (i.e. *UHT* and *control).* However, since their *prior treatment* grouping was counterbalanced with regard to their *current treatment* assignment, this orthogonal design allows us to compare any change in bodyweight from the last occasion they were measured in that study. Therefore, the bodyweight data from the current experiment (*'Enrich Post-treat.'* in Figure 3.39), and from the final measurement occasion in the previous experiment (*'UHT Post-treat. 4'* in Figure 3.39) were submitted to a repeated-measures GLM, with *measurement phase* as a within-subjects factor, and *current treatment* as a between-subjects factor. This found a significant main effect of *measurement phase,* and a significant interaction between *current treatment* and *measurement phase,* with *current treatment* non-significant (*treatment:* $F_{1,14}$=0.069, p=0.796; *measurement phase:* $F_{1,14}$=21.037, p<0.001; *measurement phase*treatment:* $F_{1,14}$=4.808, p=0.046). As Figure 3.39 indicates, bodyweight was overall higher during the current experiment, compared to when it was measured towards the end of the previous study, but the increase in the bodyweight of the *enriched treatment* group is greater than that of the *unenriched treatment* group.

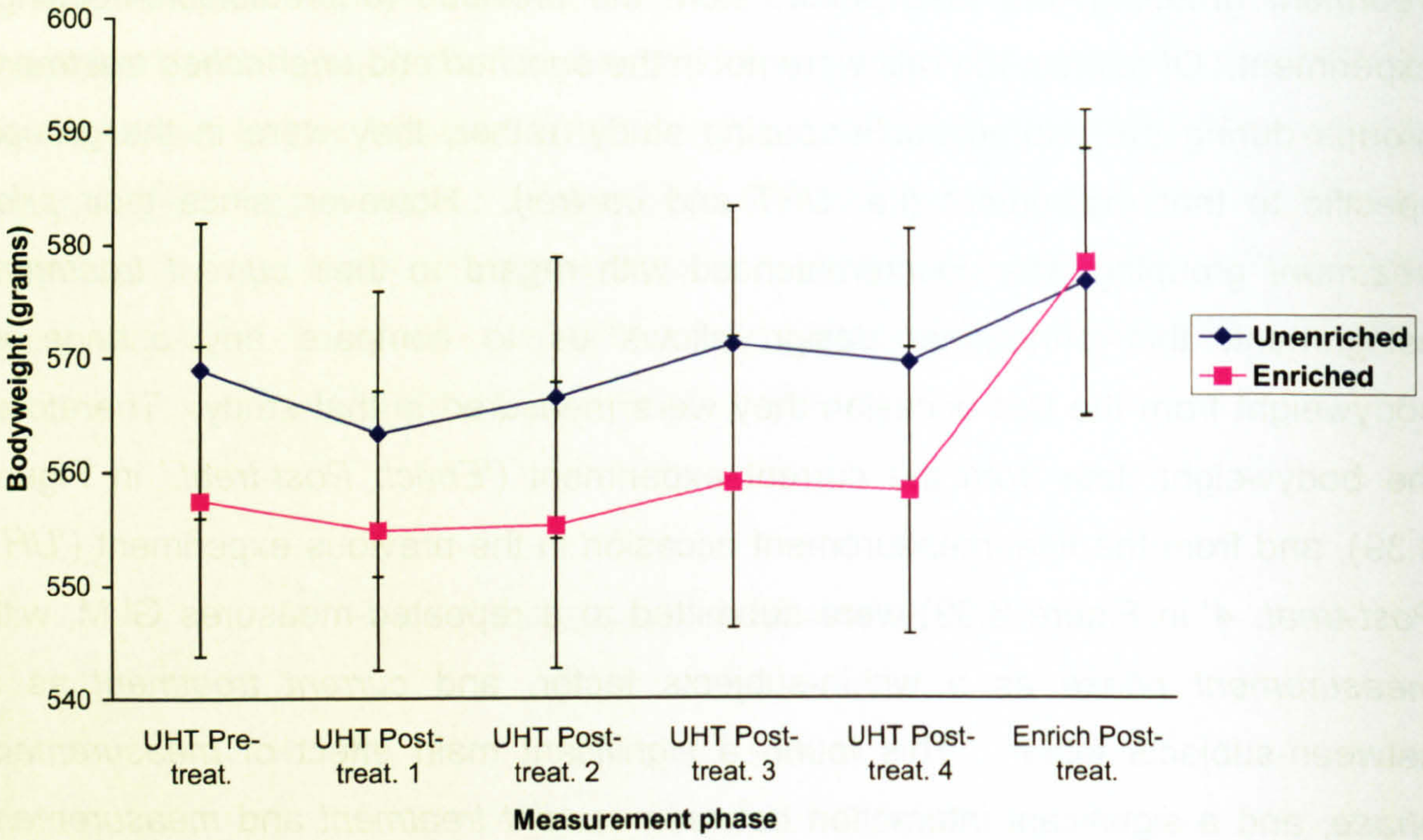

Figure 3.39 The mean rat bodyweight, by *treatment,* during different *measurement phases* (+/-1SEM). On the x-axis, *'Enrich Expt: Post-treat.'* refers to the rats' bodyweight as measured in the current (environmental enrichment) experiment, 27 days following a change in environmental enrichment. For reference, the bodyweight data from the previous (unpredictable-housing) experiment (see Figure 2.49) are also plotted (all with the prefix of *'UHT',* on the x-axis). All the data are summarised by *current treatment* group (*enriched* or *unenriched*): i.e. the data from the unpredictable-housing experiment are grouped by *future* treatment assignment (in the unpredictable-housing experiment, the rats were in treatment groups specific to that study (*UHT* and *control*); these were counter-balanced across their later assignment to the *enriched* or *unenriched* groups).

DISCUSSION

In this experiment, two treatments were employed, each designed to induce a change in rats' affective state of diverging valence: the withdrawal of 'enrichments' for the *unenriched* group was designed to induce a negative change in affect, whilst the provision of additional enrichments for the *enriched* group was designed to induce positive affective change. It was predicted that the former treatment group would be more likely to respond to probe stimuli (designed to be affectively-ambiguous) as if judging, or interpreting, them as being associated with the relatively poorer outcome of a smaller quantity of food; the latter, *enriched* group, on the other hand, were predicted to be more likely to respond to these stimuli as if

judging them to herald the *better* outcome of a larger quantity of food; i.e. the *unenriched* group were hypothesised to have a lower probability of pressing the '2-pellet' lever as a result of their treatment. We also predicted that the *unenriched* rats would be slower to press the '2-pellet' lever, and faster to press the lever associated with the relatively poorer outcome (i.e. the '1-pellet' lever), than the *enriched* group.

Our experimental hypotheses were not supported by the results: compared to the *enriched* group, the *unenriched* group were significantly *more* likely to press the '2-pellet' lever in the second *measurement phase* (i.e. when tested soon following the removal of 'enrichments') than they were in the first *measurement phase* (i.e. before the treatment), in the single-frequency probe sessions. Furthermore, they were overall faster to press the '2-pellet' lever, and slower to press the '1-pellet' lever than the *enriched* group, with no indication that this pattern changed in a manner consistent with that predicted by the hypotheses as the treatment progressed.

Elsewhere, the dual-frequency probe tests found no differences across *treatment*, nor did the progressive-ratio test of 'food motivation', nor the sucrose preference test. The analysis of the bodyweight data, however, indicated that the *enriched* group gained a significantly larger amount of weight than the *unenriched* group, at least since the last time their weight was recorded (as found by Burman et al., 2006, when they employed a similar treatment). To the extent that a change in bodyweight may be a very rough proxy of perceived stress (e.g. Broom & Johnson, 1993), including across 'environment enrichment' treatments (e.g. Young, 2003), this suggests the *unenriched* treatment may have been a relatively stressful intervention. This is perhaps not surprising: whilst the change in the 'enrichment' provision for the *enriched* group was relatively subtle (to our eyes), the change in 'enrichment' provision for the *unenriched* group seems relatively more substantial, with a loss of nesting material and shelter, which, among other things, may have otherwise provided respite from any aggressive interactions, as well as general comfort (including a potential buffer against temperature fluctuations), and cover from extraneous light (e.g. Morgan & Tromborg, 2007; Young, 2003).

So, this very brief summary of some of the key results leaves us in a position which bears some similarity to that encountered in the previous study: namely, a treatment designed to induce a negative affective state (with some, albeit tentative, evidence that it may have indeed been stressful) seems to have changed the rats' behaviour in a direction largely opposite to that predicted by our initial hypotheses. Rather than drawing our conclusions based on this relatively brief summary, though, it's worth first inspecting the broader picture: i.e. taking in elements of the rats' behaviour which fall outside the parameters our hypotheses suggest we initially inspect.

For example, during the first *measurement phase* (i.e. before the treatment had begun), the *enriched* group had a relatively poor level of accuracy in the '1-pellet' reference trials in the single-frequency probe sessions; this substantially improved, however, as the treatment progressed, to a point where, in the final *measurement phase,* the *enriched* group had a level of accuracy on both types of reference trial higher than that of the *unenriched* treatment group (e.g. see Figure 3.16 and Figure 3.17). As discussed in the last chapter, such an improvement in accuracy would, in general, be predicted if the treatment had indeed induced a more positive (or not induced a 'depressed') affective state. However, this pattern of change may have other consequences for how we interpret some of the other data from those probe sessions.

For instance, as mentioned earlier, the multi-level analysis of *lever choice* which modelled responding to *all* the probe values found that the *enriched* treatment group were less likely to press the '2-pellet' lever in the second *measurement phase,* compared to the first *measurement phase* (e.g. see Figure 3.13). This closely fits with the improvement in their accuracy with regard to the reference trials across these two *phases*, since this overall improvement was mainly achieved through the reduction of errors in the '1-pellet' reference trials (i.e. by reducing the rate the '2-pellet' lever was erroneously pressed).

It is notable, then, that in the final *measurement phase,* their level of accuracy when presented with the reference tones in the single-frequency probe sessions is even greater, with even fewer mistakes occurring in response to those reference

tones associated with one pellet of food, whereas their probability of pressing the '2-pellet' lever when presented with *all* the *probe values,* does not continue it's decline. Therefore, the *enriched* group's general performance in the task (to the extent that 'successful' performance of the task may be considered as maximising the rate of food return) improves as the treatment progresses, with increasingly less bias in the errors that they make when presented with the reference tones, yet their response to the more ambiguous *probe values* is relatively more 'optimistic' than this improvement in reference trial accuracy (and decline in bias) would, by itself, predict.

So, if we look at the data in a little more detail, it *may* be that our initial impression of the *enriched* group's more 'pessimistic-style' of responding early in the treatment phase reflected an improvement in their general effectiveness - i.e. their general accuracy – and they develop an optimistic bias in their responding to putative ambiguity despite the persistent improvement in this accuracy. Of course, this is a fairly tentative conclusion, fairly tentatively drawn, and it would be useful, in future analyses, to factor in such changes in reference accuracy (e.g. to model them as covariates), but it nonetheless illustrates the possibility that the effect of the *treatment* on the rats' responding in the operant tasks may be more complex, and subtle, than is first apparent, and may not, in fact, contradict our initial predictions.

Otherwise, as in the previous experiment, there is some evidence that the rats are still sensitive to the differential consequences of their actions: for example, there is a general, pervasive bias towards pressing the 2-pellet lever in the training sessions conducted in both the first, and final, *measurement phases* (e.g. see Figure 3.3 and Figure 3.4). This is somewhat reassuring, especially given the increasingly extensive experience the rats have with the operant tasks in question, experience which could, potentially, render their responding as more habitual, and less goal-driven (e.g. Balleine & Dickinson, 1998; Dickinson & Balleine, 1995; Dickinson et al., 1995; Phillips & Barr, 1997).

Before concluding our discussion, it's worth highlighting some interesting patterns which persist from the previous experiment. For example, the *2kHz=2pell contingency* group are still relatively more 'pessimistic' in their *lever choice*, the

effect of *contingency* across *probe value* remains fairly similar (e.g. see Figure 3.12), and there is an overall higher probability of pressing the '2-pellet' lever in the earlier *measurement phases.* In addition, the effect of the *unpredictable housing treatment* which we uncovered in the previous study seems to persist: the *UHT prior treatment* group are faster to record a response in the single-frequency probe sessions, appear to be more accurate and discriminating across *probe value* (e.g. see Figure 3.15), and are generally more 'optimistic' (i.e. more likely to press the '2-pellet' lever) than the rats previously in the *Control* group in that experiment. Interestingly, many of these differences attenuate slightly as the experiment progresses, perhaps as the effects of the *prior treatment* diminish (e.g. see Figure 3.14 and Figure 3.28).

In the next chapter we conduct our final experiment employing this particular paradigm (i.e. a 2-choice operant discrimination task), investigating the effect of food motivation on rats' responses to ambiguity, as manipulated by a pre-feeding treatment.

CHAPTER 4

THE EFFECTS OF PRE-FEEDING ON JUDGEMENTS OF AMBIGUITY IN RATS

INTRODUCTION

In Chapter 2, we discussed the possibility that certain treatments designed to induce a change in affective state, such as those employed in the previous two studies, may be correlated with, or manifested in, changes in a range of biological systems, including, for example changes in 'food motivation', hunger, and so on. We noted that the design employed by Harding et al (2004) may have been especially vulnerable to changes in the utility of the unconditional stimuli (e.g. how much subjects value access to them, or seek to avoid them, etc.), since those stimuli differed qualitatively (one food pellet vs. delivery of white noise). However, designs such as the one employed in the previous two studies (which used differing quantities of food as the unconditional stimuli) and elsewhere (e.g. Matheson et al., 2008, who used differing delays prior to food delivery), in which the unconditional stimuli differ *quantitatively*, may still be vulnerable to such confounds: for example, it is possible that a subject more 'motivated' to gain food might bias their behaviour in a manner in keeping with an 'optimistic' style of responding.

Here we explore this possibility by attempting to manipulate rats' level of 'food motivation' using a pre-feeding treatment (e.g. Babb & Crystal, 2006; Bizarro & Stolerman, 2003; Jones et al., 1990) in which subjects are given access to a fixed quantity of food (the same type of food they routinely receive as reinforcement in their operant sessions, but have rarely received outside the operant chamber[136]) before half of their probe test sessions; their remaining test sessions are administered as before: i.e. without any such pre-feeding.

[136] The only occasion the rats have received this food outside the operant chamber is the *'Time taken to eat 50 pellets of food'* test we employed in Chapter 2.

By testing the effects of a treatment designed to change rats' motivational state on their responses in the probe testing sessions, we may also gain some insight into what the rats 'know' about the operant task. As noted in the preceding two chapters, it is important, for the purposes of our hypotheses, that the experimental subjects have some operational knowledge of the consequences of their actions in the tasks we have employed; more specifically, that they have knowledge that the consequences of responding correctly to one reference stimulus (e.g. 2kHz) are different from the consequences of responding correctly to the other reference stimulus (4kHz, in this instance). For example, if we find, in the current experiment, that our pre-feeding treatment has no effect on the rats' behaviour in the probe tests, there are some grounds for concluding that their responding in those tasks has become relatively 'habitual': i.e. less goal-driven (e.g. Dickinson et al., 1995). Alternatively, if the pre-feeding treatment increases the rats' response latency (e.g. Bizarro & Stolerman, 2003), and/or decreases their accuracy in a manner insensitive to the differing quantities of food the reference trials (i.e. those in which 2kHz and 4kHz tones are presented) deliver, we would have some grounds for concluding that the animals 'know' that the task is associated with the receipt of food *per se,* but that the varying quantities of that food reward are conceptually undifferentiated. Alternatively, if we find that such changes in responding occur with a degree of asymmetry with regard to prospective reinforcement magnitude, then depending on the nature of that asymmetry, we might surmise that the rats do, in fact, have differentiated 'knowledge' of the consequences of engaging in the task, a condition which is central to the veracity of the hypotheses we have adopted in the previous two experiments.

To enable us to make a more informed decision when selecting an appropriate quantity of food to present to the rats in the pre-feeding treatment[137], we conducted brief pilot tests (not reported) employing a different set of subjects; however, to gain some additional information regarding the quantity of 'operant food' the experimental subjects in the present study would consume if given free access over a period of time roughly equal to the probe test sessions, we gave the rats a

[137] e.g. so that we were not feeding them with a quantity that was so large it risked the possibility that they wouldn't engage in the subsequent operant task at all, yet was large enough to have an appreciable effect on their 'food motivation'.

series of 'free-feeding' sessions at the end of the present study (i.e. once the probe-testing was complete), in which they were presented with *ad libitum* quantities of the food pellets.

METHOD

Overview

See Figure 4.1 for a summary of the experimental schedule.

Subjects and housing

The experimental subjects were 15 male Lister hooded rats (*Rattus norvegicus;* Harlan UK Ltd., Bicester, UK); they were previously used as *subjects* in the studies described in Chapters 2 and 3, and by the time operant training in the current experiment commenced, six weeks had passed since the termination of the (enrichment) treatment employed in the previous experiment. Since one rat (in the *4kHz=2pell contingency* group) was euthanised at the end of the last study (Chapter 3), there was one less *subject* in the current experiment.

The rats were housed in stable pairs, in cages measuring 56cm (L) x 34cm (W) x 19cm (H), with a 12:12 hour lights on:off cycle (lights off at 9am). Their homecages were cleaned on the same morning each week, and contained sawdust bedding (Lignocel), shredded paper for nesting, a red Perspex shelter (Lilico, UK) and a chew block. They were provided with *ad libitum* access to food (Eurodent Diet 22%) and water.

The rats were checked daily for health throughout the experiment.

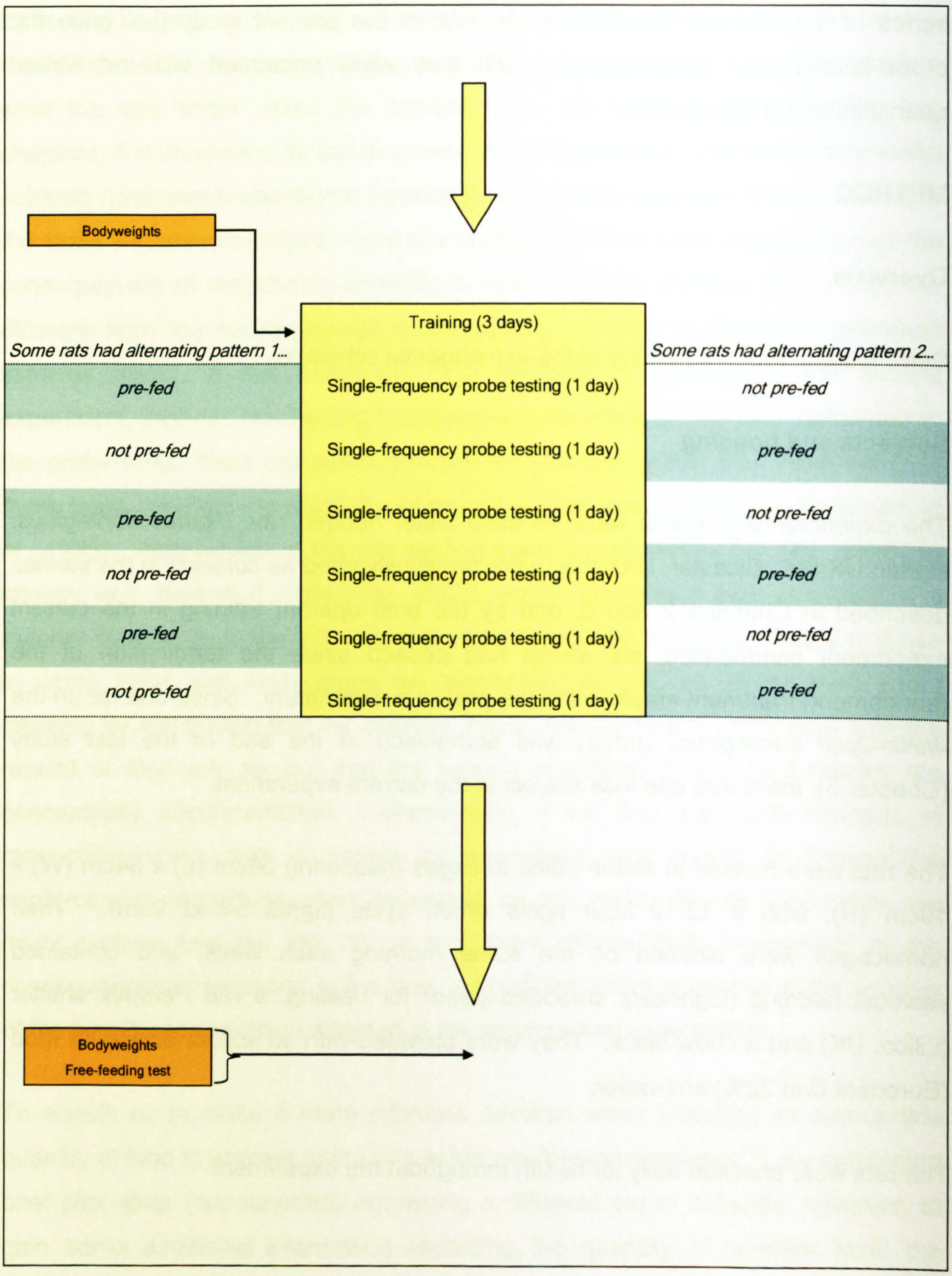

Figure 4.1 Summary of the experimental schedule.

Two-choice operant discrimination training & testing

Two-choice operant discrimination training, with differential reinforcement

The rats received one session of training per day, on the three consecutive days immediately prior to the first probe-testing session. The design of these training sessions was exactly the same as that described on p.59.

Single-frequency probe testing

The rats received one session of single-frequency probe-testing per day, over six consecutive days. The design of the single-frequency probe test sessions was exactly the same as that described on p.61.

Pre-feeding treatment

Before three of their six single-frequency probe test sessions, each rat was pre-fed 2g of the 'operant food' (i.e. Bioserv (Frenchtown, NJ, USA) Dustless Precision Pellets (45mg)) (*pre-fed treatment*); otherwise, they were not pre-fed these pellets prior to their single-frequency probe test sessions (*not pre-fed treatment*). For each rat, assignment to the *pre-fed* and *not pre-fed treatments* alternated daily over the six single-frequency probe test sessions, with the alternating order counterbalanced as far as possible with respect to prior experimental grouping; seven rats were first *pre-fed* prior to their *first* probe test session, whilst eight rats were first *pre-fed* prior to their *second* probe test session.

Prior to each single-frequency probe test session, each rat was placed, alone, in a holding cage, with *ad libitum* access to water and their usual lab chow (Eurodent Diet 22%). In the *pre-fed treatment,* a small brown bowl containing 2g of food pellets was placed centrally in the holding cage, and left for ten minutes, after which the rat was taken from the holding cage, and placed in the operant chamber (in a different experimental room) for the start of its probe testing session. The protocol for the *not pre-fed treatment* was exactly the same, except that the bowl of

food pellets was not placed in the cage; i.e. the rat simply remained in the holding cage for ten minutes prior to its probe test session.

During the *pre-feeding treatment,* all the rats ate the 2g of food pellets presented to them before their subsequent probe-testing session, bar one rat who left one pellet on one occasion (in his final *pre-fed* session).

Concurrent tests

All concurrent tests took place during the dark phase of the rats' lighting schedule.

Bodyweight

The rats were weighed, in counterbalanced order, the day before the six single-frequency probe tests began, and on each day of the free-feeding tests (and thus after the probe-testing had finished), before the free-feeding measurements began.

Free-feeding test

Each rat was placed in a test cage, with *ad libitum* access to water. A small brown bowl containing 20g of the 'operant food' (i.e. Bioserv (Frenchtown, NJ, USA) Dustless Precision Pellets (45mg)) was placed centrally in the cage[138], and left for one hour, after which the rat was returned to its homecage, and the weight of the food pellets was taken.

Each rat was tested once a day, over three consecutive days, ten days following the last single-frequency probe testing session. Each rat was tested separately, with cagemates tested simultaneously; order of testing was counterbalanced as far as possible with respect to previous experimental group assignment.

[138] Otherwise, the cage contained no food.

Data analysis

As in the last two chapters, we analysed the *lever choice* and *latency* data from the single-frequency probe sessions in multilevel models, using MLwiN. In Chapter 2 we introduced our analytical procedures in some detail, and rather than repeat that information, we will follow the format of the last chapter, and only present our main findings. The significance tests, the estimation procedures, and the basics of model specification (i.e. selection of response (y) variables, centering of continuous predictor (x) variables), remain the same as previously, as do our general model-fitting procedures. Once more, only non-reinforced trials, in which a lever press was recorded, are included in the analysis.

RESULTS

Ten of the rats completed all the trials (i.e. were never 'timed-out') in the probe test sessions which followed the *not pre-fed* treatment, whilst the same was true of eight of the rats following the *pre-fed* treatment. Averaging across *subjects*, the mean number of trials completed in the *not pre-fed* treatment was 151 (SEM: 2.64), whilst the mean number of trials completed in the *pre-fed* treatment was 142 (SEM: 6.09).[139]

Single-frequency test sessions: lever choice

Figure 4.2 and Figure 4.3 chart the mean percentage of presses on the lever associated with 2 pellets of food (as opposed to the lever associated with 1 pellet of food) across kHz, by *pre-feeding treatment* for the *2kHz=2pell* and *4kHz=2pell contingency* groups, respectively.

[139] Note, the *maximum* number of possible trials in a session (i.e. when it was not 'timed-out) was 156.

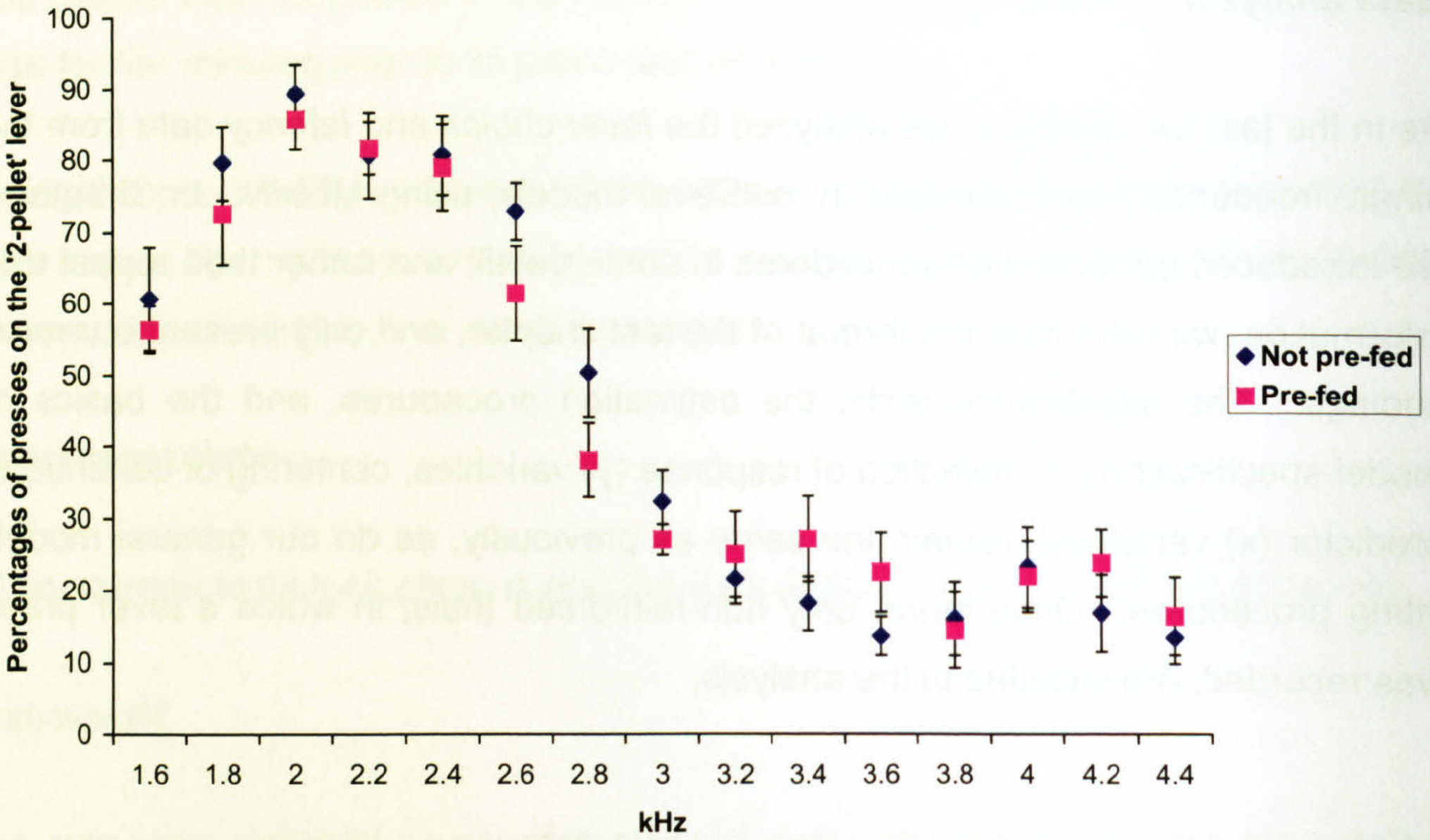

Figure 4.2 *2kHz=2pell.* The mean percentage of lever presses made on the '2-pellet' lever (as opposed to the '1-pellet' lever) for the *2kHz=2pell contingency* group, by *pre-feeding treatment*, across kHz (+/- 1 SEM. N.B. the means and SEM are derived from data summarised at the *subject*-level; in addition, the data pertaining to the 'reference tones' (i.e. 2kHz and 4kHz) are taken from the non-reinforced trials only).

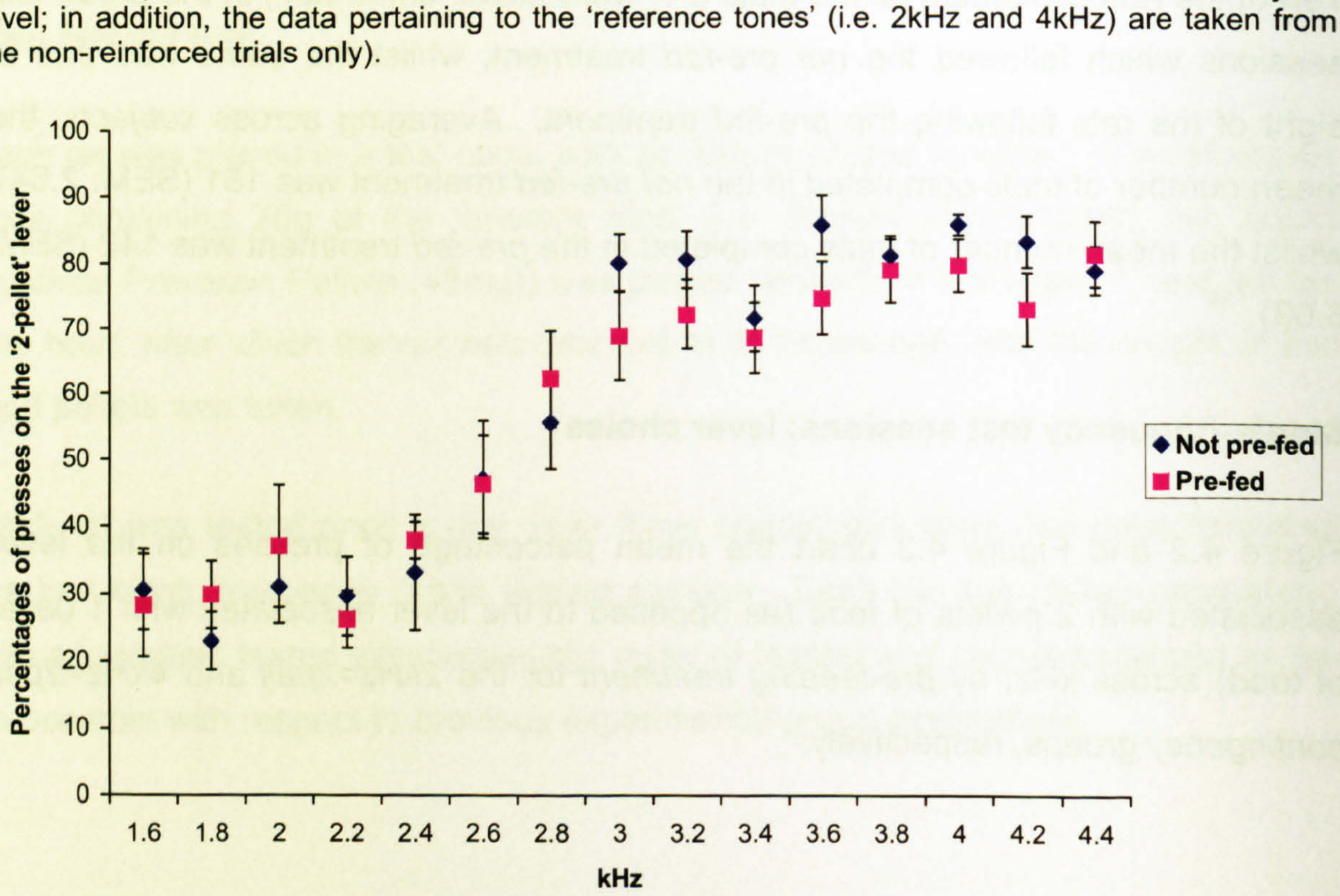

Figure 4.3 *4kHz=2pell.* As Figure 4.2, but for the *4kHz=2pell contingency* group.

Multi-level multiple logistic regression in MLwiN

The dataset was once more defined as having two hierarchical levels, with *trial* (n=7,565) at the lowest level of the hierarchy (Level 1), and *subject* (n=15) at the next, higher, level of the hierarchy (Level 2; i.e. *trial* was nested within *subject*).

A 'random slope' model was again fitted, up to a cubic term for *probe value*[140], with the intercept, and the linear and quadratic *probe value* terms, allowed to vary at the *subject*-level. As Table 4.1 illustrates, this process revealed a significant difference between *subjects* both in the overall probability of pressing the '2-pellet' lever, and in the probability of *lever choice* across *probe value.* There was also a significant fixed (overall) effect of *probe value*, indicating that the polynomial model was a good fit to the data, and that there was a significantly greater probability of pressing the '2-pellet' lever as the linear *probe value* term increased: i.e. as it approached the end of the scale where the '2-pellet' reference tone is located.

The categorical predictor variables of main interest (*contingency, current (pre-feeding) treatment, prior (enrichment / unpredictable housing) treatment*), were then systematically added to the model as fixed effects.

As Table 2.7 shows, *Contingency*, as a main effect[141], was not a significant term, but the two-way interactions between *contingency* and *probe value* were, indicating that the probability of pressing the 2-pellet lever across the standardised scale of *probe value* differed between the two *contingency* groups.

---

[140] We used the standardised scale of *probe value,* as described on p.71

[141] The *2kHz=2pell* group was the reference category, assigned a value of *'0'*, whilst the *4kHz=2pell* group was assigned a value of *'1'*.

Parameter		Coefficient estimate (with SE)	Wald (χ^2)	Df	*P*
Intercept	Random at *subject* level	0.168 (0.064)	6.829	1	0.009 **
Probe value	Random at *subject* level	0.647 (0.263)	6.044	1	0.014 *
	Fixed	2.492 (0.219)	129.442	1	<0.001 **
Probe value	Random at *subject* level	0.697 (0.282)	6.098	1	0.014 *
	Fixed	4.639 (0.263)	310.289	1	<0.001 **
(Probe value)^2	Random at *subject* level	0.957 (0.471)	4.121	1	0.042 *
	Fixed	-0.698 (0.294)	5.657	1	0.017 *
(Probe value)^3	Fixed	-5.926 (0.366)	261.443	1	<0.001 **

Table 4.1 The coefficient estimates, with standard error, together with Wald test statistics and their significance, of fixed and random parts of various models fitted in gradual increments up to a polynomial 'random slope' model (as specified in Equation 0.92, in the Appendix).

Parameter		Coefficient estimate (with SE)	Wald (x2)	Df	*P*
Contingency	Fixed	0.170 (0.198)	0.738	1	0.390
*Contingency*Probe value*	Fixed	-1.994 (0.504)	15.632	1	<0.001 **
Contingency(Probe value)*^{2}	Fixed	-1.341 (0.517)	6.729	1	<0.009 **
Contingency(Probe value)*^{3}	Fixed	4.034 (0.800)	25.403	1	<0.001 **

Table 4.2 The coefficient estimates, with standard error, of selected fixed parts of the model, together with Wald test statistics and their significance (from the model specified in Equation 0.93, in the Appendix).

As Table 4.3 indicates, *treatment* did not have a significant main effect[142], indicating there was no overall difference in the probability of pressing the '2-pellet' lever depending on whether *subjects* were pre-fed or not prior to their test session. When two-way interactions of *treatment* with the *probe value* terms were added, however, the interaction with the linear *probe value* term was highly significant, whilst the interaction of *treatment* with the cubic *probe value* term bordered significance at the 0.05 level. The valence of the coefficient of the linear *probe value* term indicates that the overall linear trend is less positive (i.e. less steep) for the rats in the *pre-fed treatment* group (e.g. p.65, Aiken & West, 1991); this is illustrated in Figure 4.4, which plots the predicted effect of *pre-feeding* on *lever choice,* across *probe value,* contrasted against the *not pre-fed treatment,* which is held constant at a probability of 0.5.

Otherwise, a two-way interaction between *treatment* and *contingency* was not a significant addition to the model, nor were three-way interactions between *treatment, contingency* and *probe value*.

Parameter		Coefficient estimate (with SE)	Wald (x2)	Df	*P*
Treatment	Fixed	-0.090 (0.055)	2.684	1	0.101
*Treatment*Probe value*	Fixed	-0.918 (0.278)	10.901	1	<0.001 **
Treatment(Probe value)*^{2}	Fixed	0.141 (0.278)	0.256	1	0.613
Treatment(Probe value)*^{3}	Fixed	1.210 (0.620)	3.806	1	0.051

Table 4.3 The coefficient estimates, with standard error, of selected fixed parts of the model, together with Wald test statistics and their significance (from the model specified in Equation 0.94, in the Appendix).

[142] The *Not pre-fed treatment* group was the reference category, assigned a value of *'0'*, whilst the *Pre-fed treatment* group was assigned a value of *'1'*.

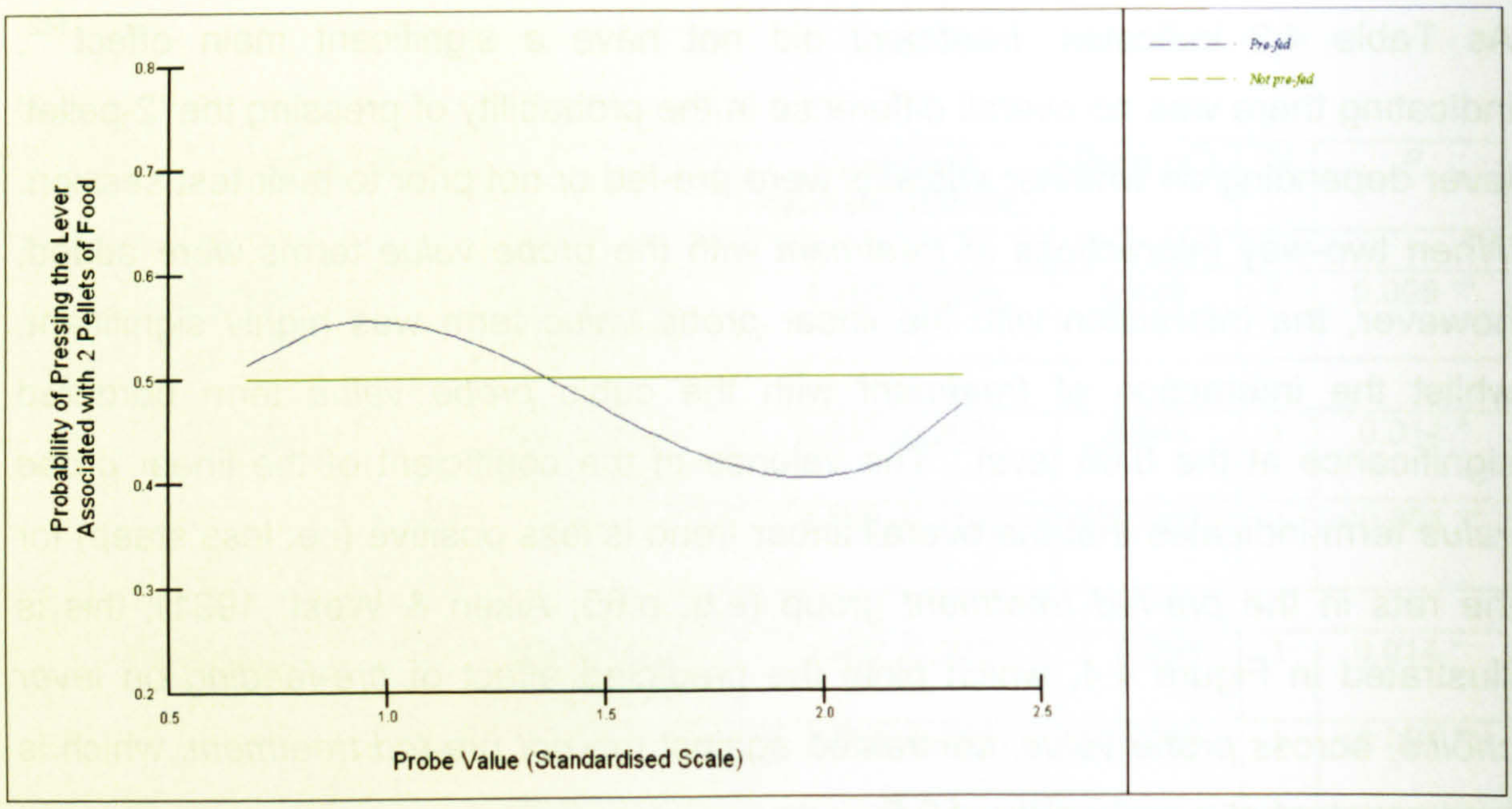

Figure 4.4 The predicted effect of the *pre-feeding* on the probability of pressing the '2-pellet' lever, across *probe value* (a standardised scale, with *'1'* corresponding to the reference stimulus associated with 1 pellet of food, and *'2'* corresponding to the reference stimulus associated with 2 pellets of food. Prediction equation derived from the model specified in Equation 0.94). Here, the *pre-fed treatment* group is compared against the reference category of the *Not pre-fed* group, which has a probability held constant at 0.5.

The *prior treatment* terms were then systematically added to the model. Firstly, the treatment group the rats were assigned to in the first experiment in which they were *subjects* (described in Chapter 2) was added to the model as a main effect[143]. As Table 4.4 indicates, this term was not a significant addition to the model, and nor were the various interactions we explored with *probe value, pre-feeding,* and *contingency* (values not reported).

In contrast, as Table 4.4 also shows, the other *prior treatment* term, namely that pertaining to *treatment* assignment in the *environmental enrichment* study (see Chapter 3), did have a significant main effect[144], indicating that those rats previously in the *enriched treatment* group had a higher probability of pressing the

---

[143] The *Control* group were the reference category, assigned a value of *'0'*, whilst the *unpredictable housing treatment (UHT)* group was assigned a value of *'1'*.

[144] The *non-enriched treatment* group were the reference category, assigned a value of *'0'*, whilst the *enriched treatment* group was assigned a value of *'1'*.

lever associated with 2 pellets of food, compared to those rats previously in the *non-enriched* group. Otherwise, various interactions of *enrichment treatment group* with *probe value, pre-feeding, contingency,* and *unpredictable-housing treatment* group were explored, in a variety of models, but all were non-significant (values not reported).

Parameter	Coefficient estimate (with SE)	Wald ($\chi 2$)	Df	*P*
Unpredictable-housing treatment group	Fixed	0.159 (0.197)	0.652	1	0.419
Enrichment treatment group	Fixed	0.437 (0.168)	6.794	1	0.009 **

Table 4.4 The coefficient estimates, with standard error, of selected fixed parts of the model, together with Wald test statistics and their significance (from the models specified in Equation 0.95 and Equation 0.96, in the Appendix).

Finally, a few other terms of interest were investigated in the model. As Table 4.5 indicates, *counterbalanced pre-feeding group*[145] did not have a significant main effect, nor did the mean weight of food pellets eaten during the later free-feeding sessions; however, *session number* (i.e. 1 to 6), had a highly-significant main effect, revealing a significantly lower probability of pressing the lever associated with 2 pellets of food as *session number* progressed: i.e. the rats were more likely, overall, to press the '1-pellet' lever, in the later sessions. It's worth noting that even though the alternation of the *pre-feeding treatment* across the six sessions was counterbalanced as far as possible, there was a loss of orthogonality in the experimental design due to the odd number of *subjects* (n=15), therefore more of the *pre-fed* sessions appeared later than the *non-pre-fed* sessions; however, since *pre-feeding* did not exert a significant main effect on *lever choice*, it seems unlikely this is responsible for the significant main effect of *session number.* Furthermore, as Table 4.5 indicates, the *latency* to record a lever response was not a significant main effect, indicating there was no overall difference in *latency* depending on

[145] As outlined in the Method section, for each *subject,* the test sessions before which they were *pre-fed* alternated with those before which they were *not pre-fed*, so that each *subject* had three alternating sessions of each, once a day, across six days. This alternating order was counterbalanced, as far as possible, with respect to whether the rats were pre-fed in their first test session, or not.

which of the two levers were pressed. Of course, *latency* may be an important factor in other respects, and we explore this further when we model it as the response (y) variable, below.

Finally, whilst *bodyweight* (as measured the day before probe-testing began), as a main effect, was significant, indicating that heavier rats were overall more likely to press the '2-pellet' lever (Wald=9.559, 1 d.f., p=0.002), the size of the regression coefficient was very small (coefficient estimate (SE) = 0.006 (0.002)). Since the *bodyweight* variable was actually measured in quite small units (grammes, rather than kilogrammes, for example) the size of the coefficient estimate suggests the effect was unlikely to be of much biological interest (e.g. Nakagawa & Cuthill, 2007). A similar situation was encountered when adding an interaction of *bodyweight* with *probe value:* the interaction featuring the linear *probe value* term was significant (Wald=6.392, 1 d.f., p=0.011; the interaction with the other *probe value* terms was not significant) indicating that the resulting response curve was less steep for heavier rats (i.e. they were less discriminating across *probe value*), but the coefficient estimate was again relatively small (-0.012 (0.005)). Otherwise, there were no significant interactions between *bodyweight* and *pre-feeding treatment,* either when *probe value* was included in the interaction term, or not, indicating that heavier rats were not more, or less, likely to be effected by the *pre-feeding treatment* with regard to their *lever choice.*

Table 0.6 and Figure 0.51, in the Appendix, present the results of a Hosmer-Lemeshow goodness-of-fit test for the final fitted model, which proved satisfactory.

Parameter		Coefficient estimate (with SE)	Wald (χ2)	Df	*P*
Pre-feeding counterbalanced group	Fixed	0.251 (0.156)	2.599	1	0.107
Mean weight food pellets eaten in 1 hour free-feed	Fixed	-0.027 (0.049)	0.298	1	0.585
Session number	Fixed	-0.064 (0.016)	15.199	1	<0.001 **
Latency	Fixed	-0.002 (0.002)	1.780	1	0.182

Table 4.5 The coefficient estimates, with standard error, of selected fixed parts of the model, together with Wald test statistics and their significance (the values relating to *Session number* are derived from the model specified in Equation 0.97, whilst all other terms are derived from the model specified in Equation 0.96, with the addition of that particular term; all in the Appendix).

Reference trial accuracy

For each *subject*, the mean percentage of 'correct' responses in *all* the reference trials (i.e. reinforced or not – therefore modelling a larger number of such trials) in the single-frequency probe sessions was pooled within each *treatment,* and submitted to exploratory analyses, using separate repeated-measures GLMs for each trial type (i.e. all reference trials, '1-pellet' reference trials, '2-pellet' reference trials), with *treatment* as a within-subjects factor, and *contingency* as a between-subjects factor. These found a significant main effect of *treatment* when modelling *all* reference trials, and also when modelling just the '2-pellet' reference trials; all other main effects and interactions were non-significant (all reference trials - *treatment:* $F_{1,13}$=4.906, p=0.045; *contingency:* $F_{1,13}$=3.133, p=0.1; *treatment*contingency*: $F_{1,13}$=0.807, p=0.385; '2-pellet' reference trials (square-transformed) - *treatment:* $F_{1,13}$=4.763, p=0.048; *contingency:* $F_{1,13}$=2.025, p=0.178; *treatment*contingency*: $F_{1,13}$=0.472, p=0.504; '1-pellet' reference trials - *treatment:* $F_{1,13}$=3.125, p=0.101; *contingency:* $F_{1,13}$=2.007, p=0.180; *treatment*contingency*: $F_{1,13}$=0.597, p=0.453). Figure 4.5 plots this data, illustrating that when in the *pre-*

fed treatment, the rats were, on average, less accurate when presented with the reference tones.

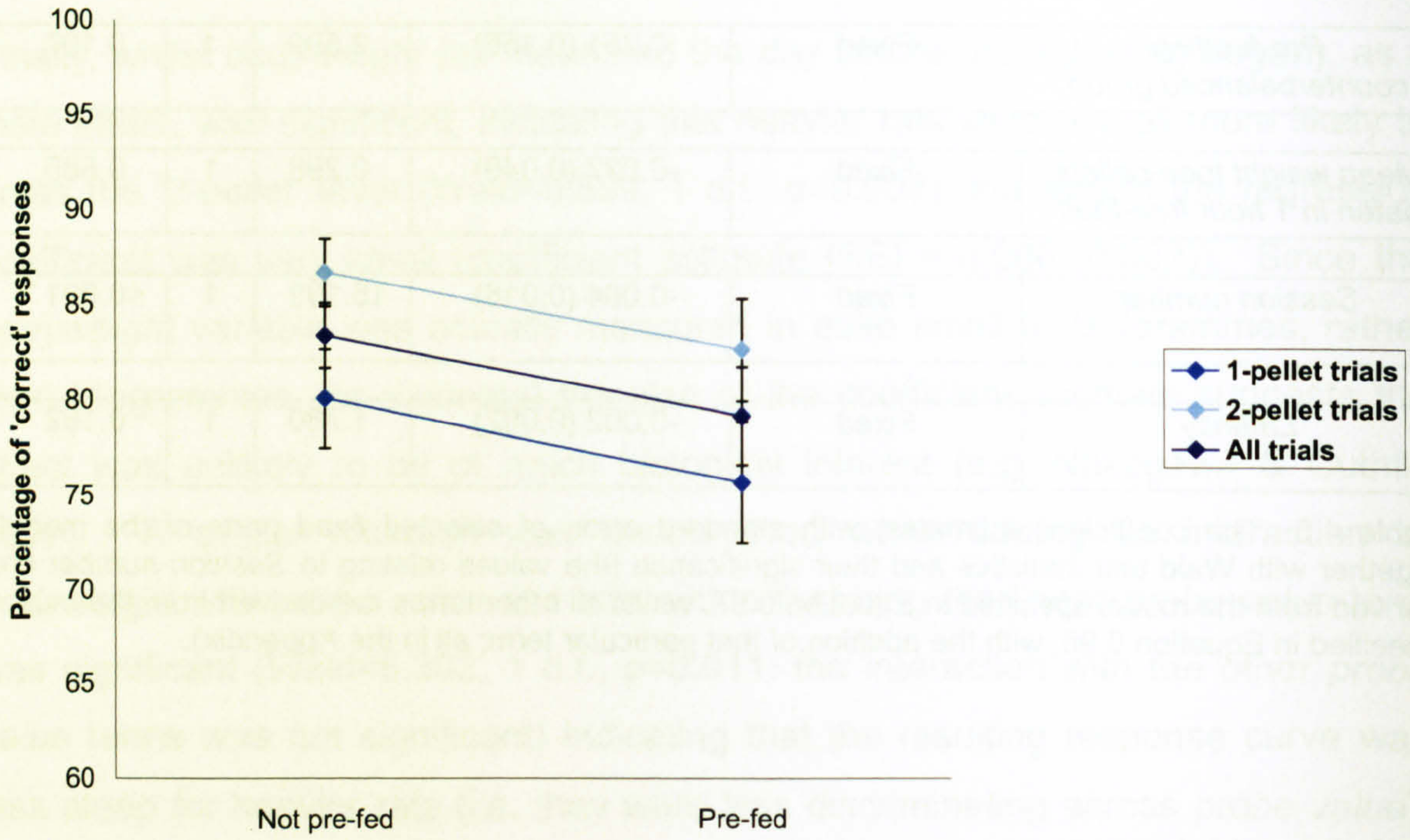

Figure 4.5 The mean percentage of 'correct' responses in the reference trials (i.e. those in which 2kHz or 4kHz were presented, including both reinforced, and non-reinforced trials) in the single-frequency probe sessions, summarised by trial type (i.e. those tones associated with one pellet of food; those associated with two pellets of food; and all reference trials combined) across *treatment* (+/- 1SEM).

For reference, Figure 4.6 and Figure 4.7 plot the change in reference trial accuracy as each single-frequency probe session progressed, averaged across all reference trials (Figure 4.6), and averaged across each type of reference trial (Figure 4.7)[146], respectively. Figure 4.6 indicates that the advantage in overall reference trial accuracy the above analyses revealed for the *not pre-fed* group is especially apparent in the first half of the probe-testing sessions, after which the overall accuracy of the two *treatment* groups becomes more similar, as the session progresses. Otherwise, it also reveals, for both groups, an initial improvement in accuracy over the first few trials of the session.

[146] We plot these on different charts so that the resulting plots are a little clearer.

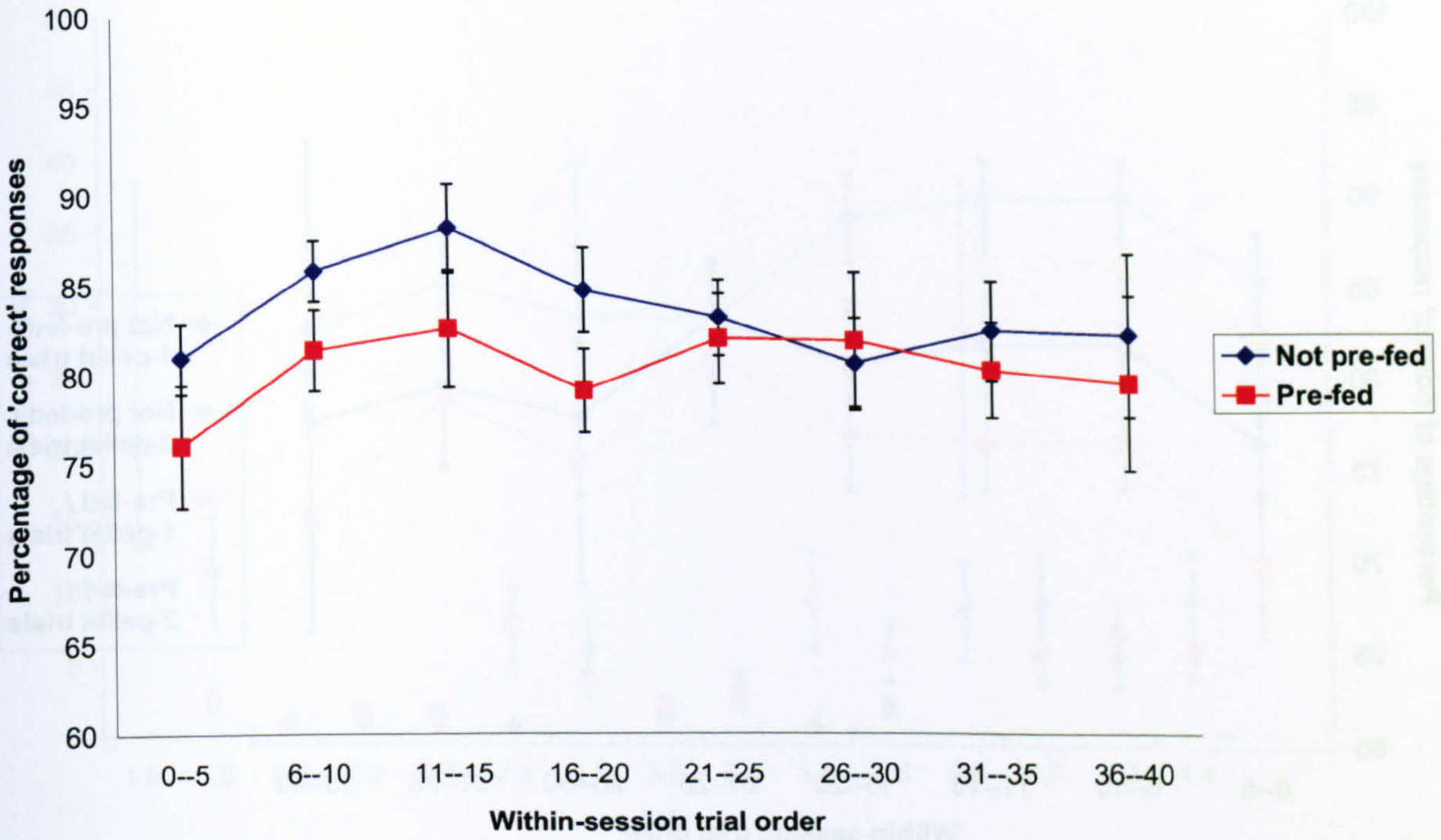

Figure 4.6 The mean percentage of 'correct' responses in the reference trials (i.e. those in which 2kHz or 4kHz were presented, including both reinforced, and non-reinforced trials) in the single-frequency probe sessions, across *treatment* (+/- 1SEM). The data are grouped into bins, each of which refers to the order in which those specific trials occurred within that session (so, bin 0-5, for example, refers to the first five reference trials in which 2kHz was presented, and the first five reference trials in which 4kHz was presented; i.e. the bin refers to ten trials). Each datapoint is only summarised from those rats for whom 20 or more responses were recorded in that particular bin; since some rats were 'timed-out' before completing all the scheduled trials, not all completed the maximum of 30 reference trials in each bin (i.e. 10 trials x 3 sessions for each *subject* in each *treatment*).

Figure 4.7 reveals that, for both *treatment* groups, an initial bias towards greater accuracy with regard to the '2-pellet' reference trials at the start of the probe-testing sessions attenuates (e.g. in trials 21-25), but then appears to be somewhat re-established towards the end of each session; this attenuation occurs a little earlier for the *pre-fed* rats (i.e. between trials 11-15 & trials 16-20) than for those *not pre-fed* (i.e. between trials 16-20 & 21-25). In the first phase of this '2-pellet' bias (i.e. up until around trials 21-25), the *non pre-fed* group appear to have overall better accuracy (as indicated in Figure 4.6), but this seems less the case in the second phase of '2-pellet' trial bias, towards the end of the session (i.e. from trials 26-30 onwards).

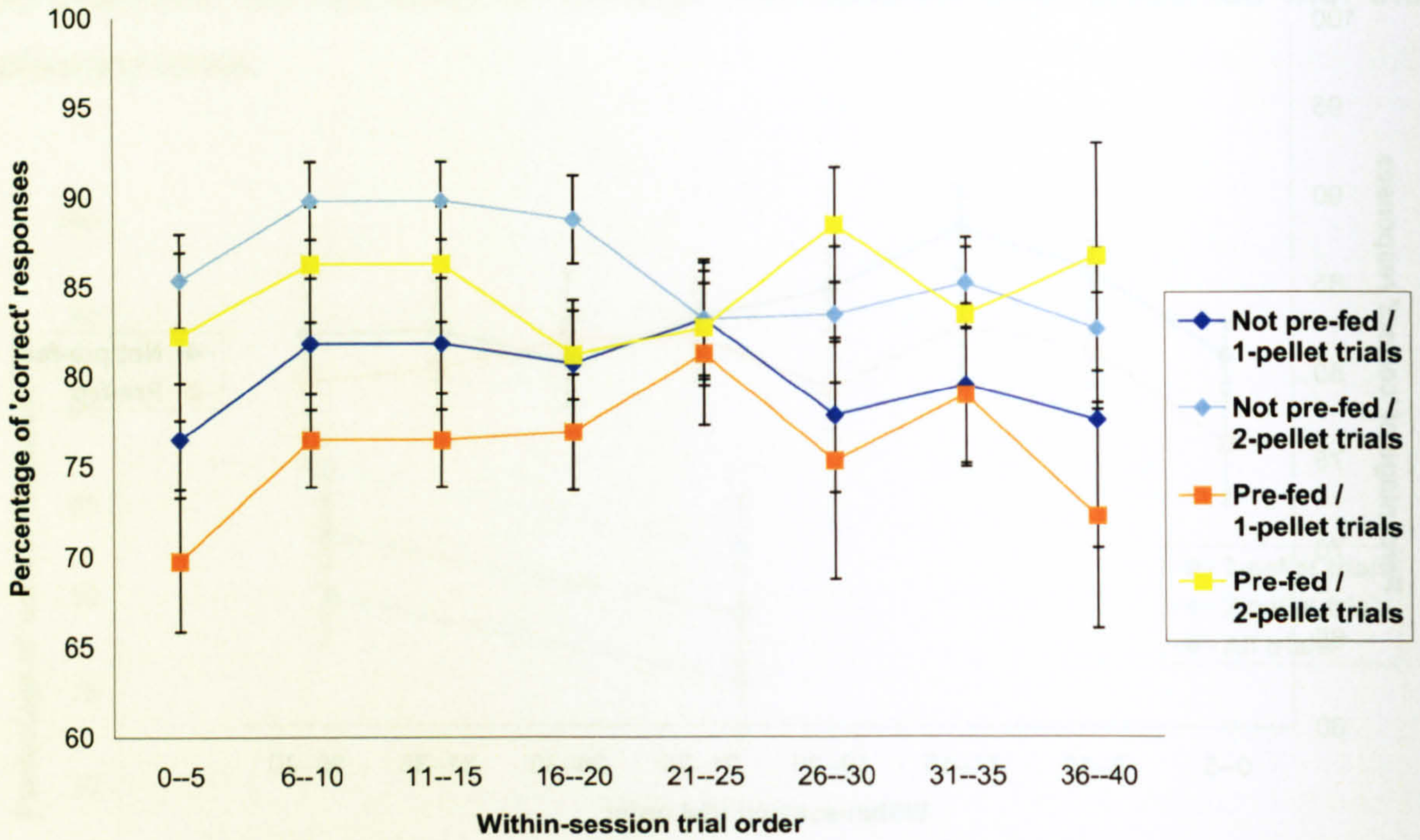

Figure 4.7 The mean percentage of 'correct' responses in the reference trials (i.e. those in which 2kHz or 4kHz were presented, including both reinforced, and non-reinforced trials) in the single-frequency probe sessions, across *treatment*, summarised by trial type (i.e. whether that particular tone was associated with one, or two, pellets of food; +/- 1SEM). The data are grouped into bins, each of which refers to the order in which those specific trials occurred within that session. Each datapoint is only summarised from those rats for whom 10 or more responses were recorded in that particular bin; since some rats were 'timed-out' before completing all the scheduled trials, not all completed the maximum of 15 reference trials, of each trial type, in each bin.

Note: we also conducted some exploratory analyses on the *latency* to respond to the reference trials (specifically) in the single-frequency probe sessions; these are presented at the end of the following section.

Single-frequency test sessions: latency

Figure 4.8 and Figure 4.9 plot the mean latency to press the lever associated with 2 pellets of food by *pre-feeding treatment*, across kHz, for the *2kHz=2pell* and *4kHz=2pell contingency* groups, respectively, whilst Figure 4.10 and Figure 4.11 plot the mean latency to press the lever associated with 1 pellet of food by *pre-feeding treatment*, across kHz, for the *2kHz=2pell* and *4kHz=2pell contingency* groups, respectively.

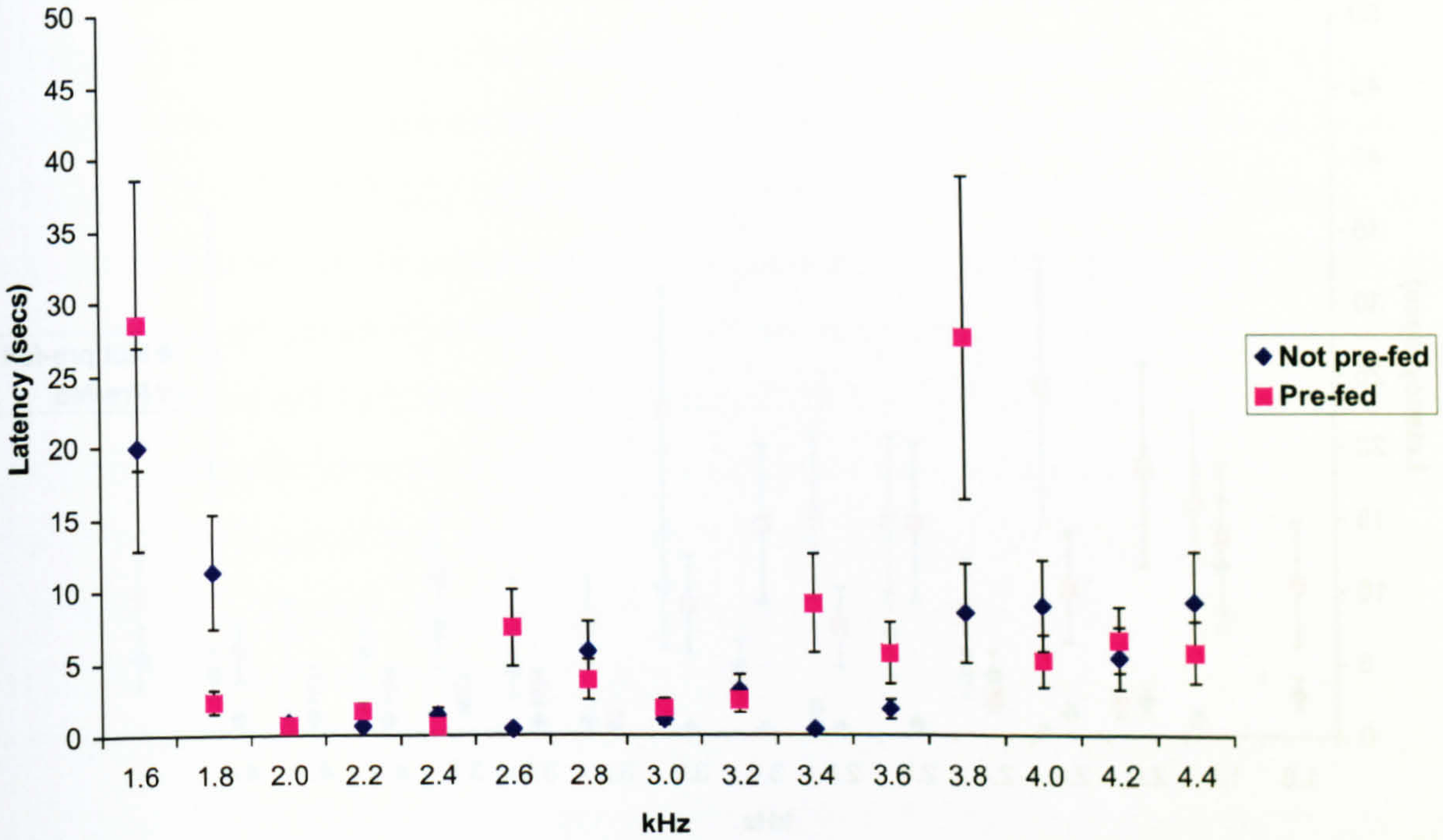

Figure 4.8 '2-pellet' lever latency: *2kHz=2pell*. The mean latency, in seconds, for the *2kHz=2pell contingency* group to press the lever associated with 2 pellets of food, by *pre-feeding treatment* group, across kHz (+/- 1 SEM. N.B. the means and SEM are derived from data summarised at the *subject*-level (note: if, in one of the *measurement phases*, a *subject* never pressed this particular lever following one of the kHz tones, that *subject* did not contribute to the corresponding summary datapoint[147]); in addition, the data pertaining to the 'reference tones' (i.e. 2kHz and 4kHz) are taken from the non-reinforced trials only). Note the scale on the y-axis, which differs from the equivalent charts in the last two chapters.

[147] An alternative would have been to enter a maximum 'timed-out' latency, but since each trial terminated following a lever press (rather than being timed-out), it is not clear what value any such maximum latency would take.

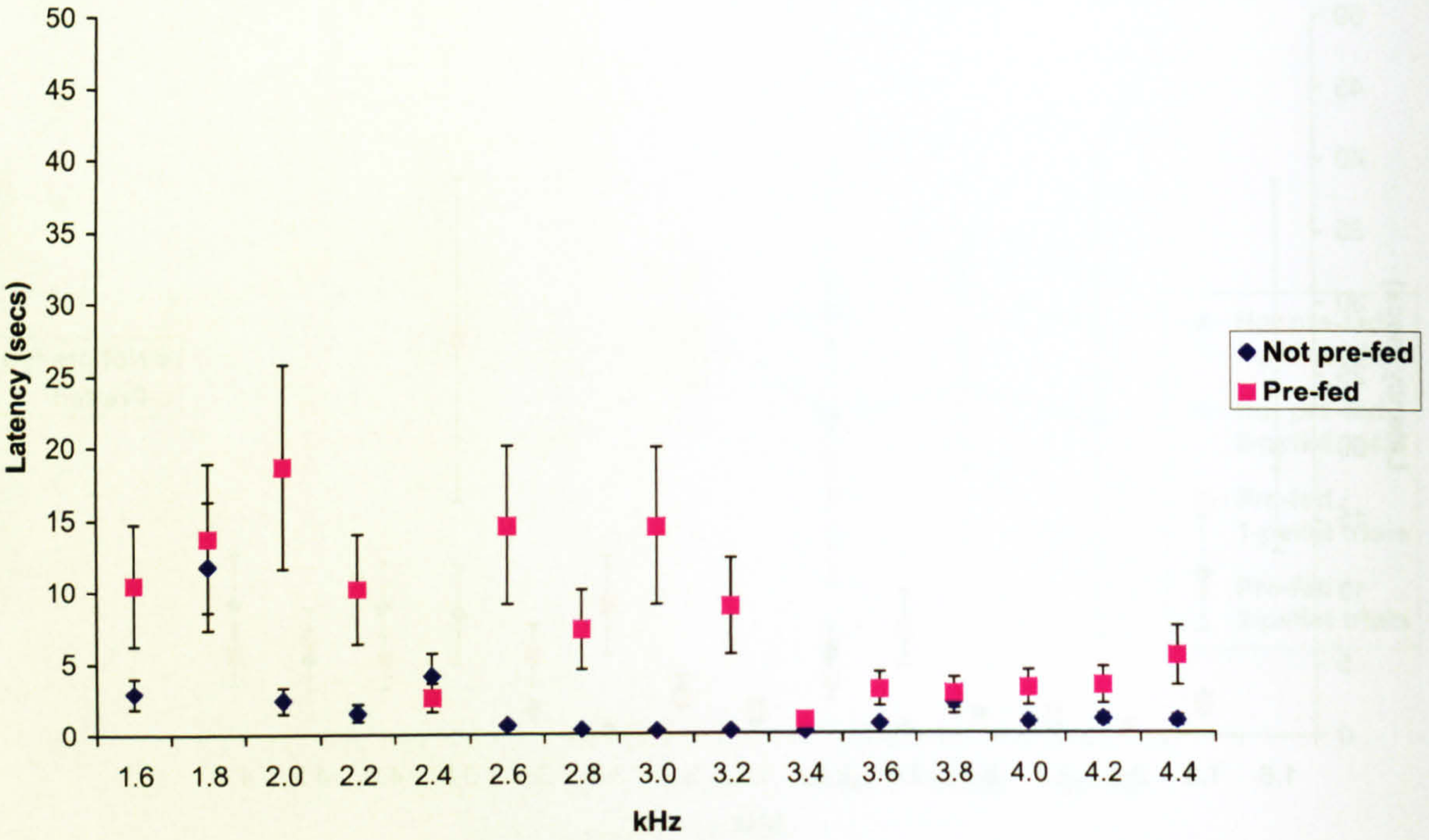

Figure 4.9 '2-pellet' lever latency: *4kHz=2pell*. As Figure 4.8, but for the *4kHz=2pell contingency* group.

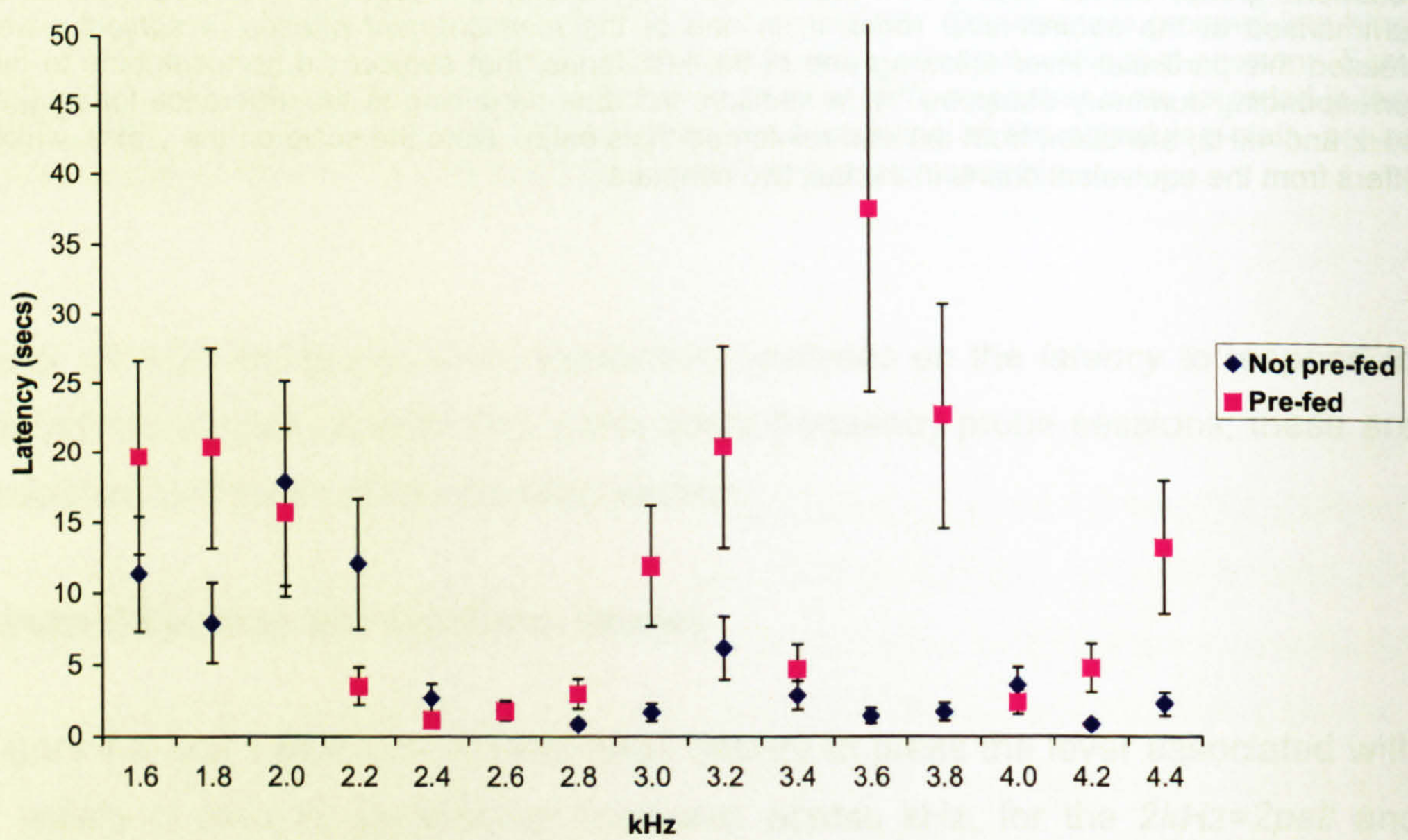

Figure 4.10 '1-pellet' lever latency: *2kHz=2pell*. As Figure 4.8, but for presses made on the '1-pellet' lever.

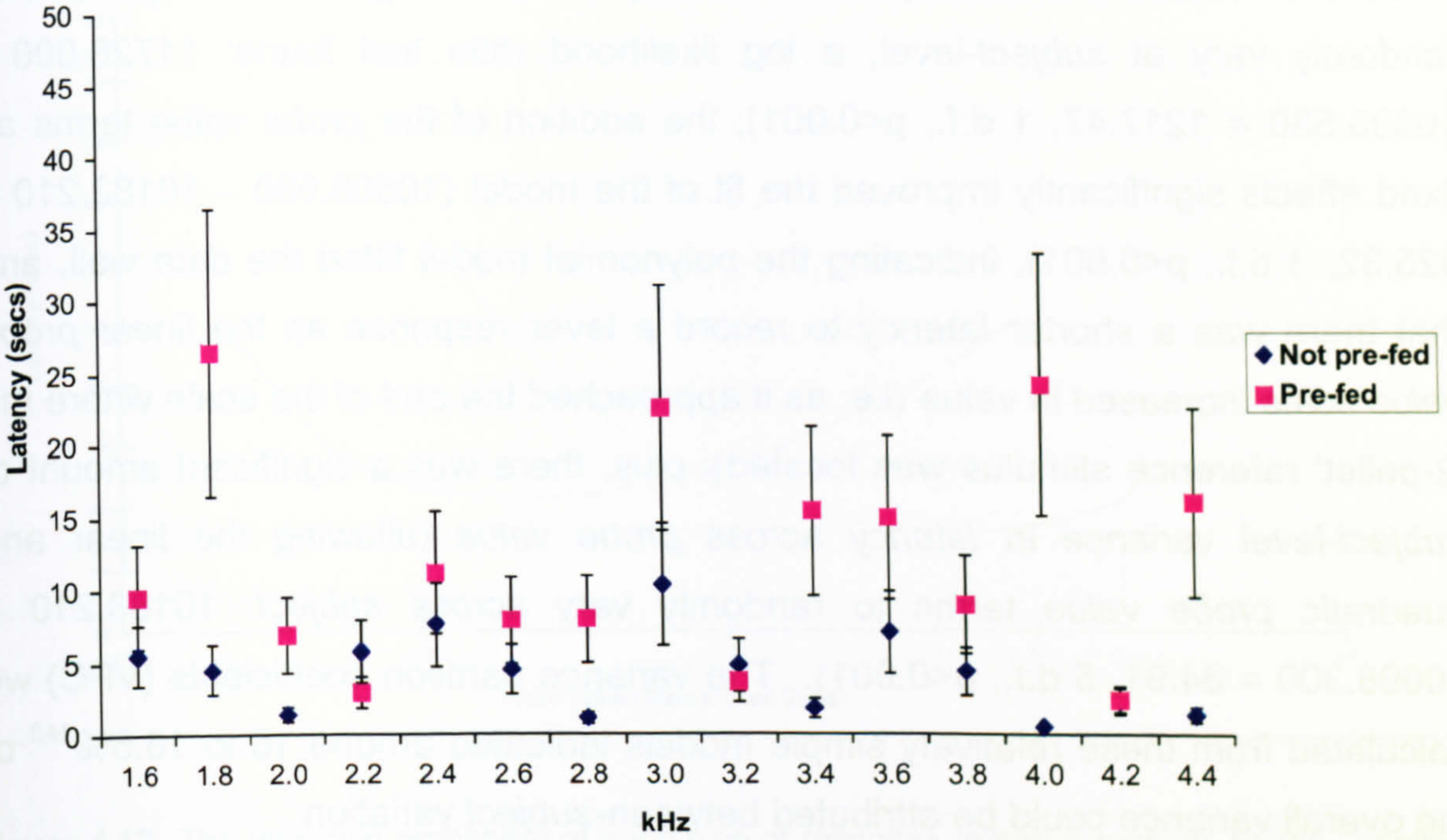

Figure 4.11 '1-pellet' lever latency: *4kHz=2pell*. As Figure 4.10, but for the *4kHz=2pell contingency* group.

Multi-level general linear model in MLwiN

As in the equivalent analysis in the last chapter, in order to better meet the assumptions of the model, we employed a negative reciprocal root transformation having first rounded the data into slightly larger 'bins' of 0.1 seconds, with a minimum *latency* of 0.2 seconds. See later for a further examination of whether the data met the assumptions of the model.

The hierarchy of the dataset was defined in the same way as in the multilevel analysis of *lever choice*, above: i.e. with *trial* (n=7,565) at Level 1, nested within *subject* (n=15) at Level 2.

The regression equation was once more refined, in increments, up to a random slope model, with the linear and quadratic *probe value* terms allowed to vary at the *subject*-level, and the cubic *probe value* term remaining fixed (Equation 0.98, in the Appendix). This process indicated the following: there was a highly-significant

amount of variation in *latency* between *subjects* (allowing the intercept term to randomly vary at *subject*-level, a log likelihood ratio test found: 11726.000 - 10508.530 = 1217.47, 1 d.f., $p<0.001$); the addition of the *probe value* terms as fixed effects significantly improved the fit of the model (10508.530 – 10183.210 = 325.32, 3 d.f., $p<0.001$), indicating the polynomial model fitted the data well, and that there was a shorter latency to record a lever response as the linear *probe value* scale increased in value (i.e. as it approached the end of the scale where the '2-pellet' reference stimulus was located); plus, there was a significant amount of *subject*-level variance in *latency* across *probe value* (allowing the linear and quadratic *probe value* terms to randomly vary across *subject:* 10183.210 – 10098.300 = 84.91, 5 d.f., $p<0.001$). The variance partition coefficients (VPC) we calculated from these relatively simple models indicated around 16 to 16.5%[148] of the overall variance could be attributed between-*subject* variation.

Whilst the valence of the coefficient of *Lever choice*[149] (i.e. whether the lever associated with 2 pellets of food was pressed or not), suggested *subjects* were overall quicker to press the '2-pellet' lever, it did not have a significant main effect in the model (10098.300 – 10094.940 = 3.36, 1 d.f., $p=0.067$), as indicated when *latency* was introduced as a predictor (x) variable in the multilevel analysis of *lever choice*, above. However, the addition of *lever choice*probe value* interaction terms did substantially improve the model's fit to the data (10094.940 – 9409.253 = 903.58, 3 d.f., $p<0.001$; Equation 0.99, in the Appendix), indicating that the *latency* to press each lever differed across *probe value*, as Figure 4.12 clearly indicates.

---

[148] Calculated first without, and then with, the *probe value* terms, respectively. These values are slightly lower than those calculated in the equivalent analysis in the last chapter (we found 17%, and 17.5%, respectively, in that analysis), which in turn, were slightly lower that those calculated in Chapter 2: 17.4% and 18%).

[149] A press on the '1-pellet' lever was assigned the reference category, with a value of *'0'*, whilst a press on the '2-pellet' lever was assigned a value of *'1'*.

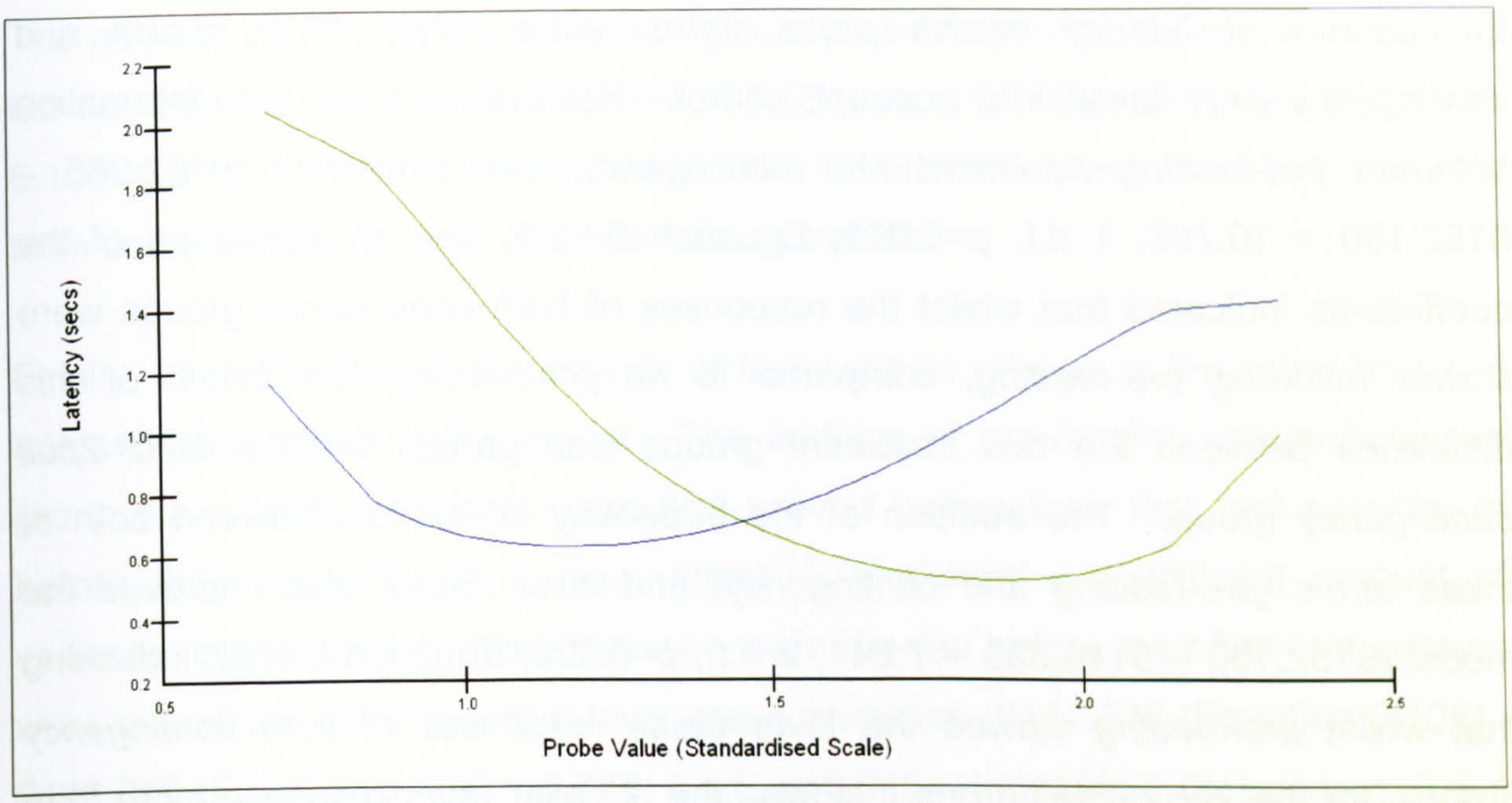

Figure 4.12 The predicted probability of each type of response (green = presses on the '2-pellet' lever; blue = presses on the '1-pellet' lever), across *probe value* (these predictions were generated from the model specified in Equation 0.99, in the Appendix).

As in the equivalent analysis in the last chapter, *Contingency* was not a significant main effect[150] (9409.253 – 9409.222 = 0.031, 1 d.f., p=0.860), but fitting *contingency* up to a three-way interaction with *lever choice* and *probe value*, did improve the fit of the model, and so these terms remained (9369.465 – 9330.703 = 38.762, 3 d.f., p<0.001; Equation 0.100).

The addition of *pre-feeding treatment*, as a main effect[151], made a very significant improvement to the fit of the model (9330.703 – 9162.965 = 167.738, 1 d.f., p<0.001; Equation 0.101, in the Appendix), indicating that when *subjects* had been *pre-fed* prior to their test session, they were considerably slower to make a lever press response. Interactions of *pre-feeding treatment* with *probe value*, *lever choice* and *contingency* were explored, in a variety of models; none of the interactions with *probe value* were significant, indicating there was no difference in

---

[150] The *2kHz=2pell contingency* group was assigned the reference category, with a value of *'0'*, whilst the *4kHz=2pell contingency* group was assigned a value of *'1'*.

[151] *No pre-feeding* was assigned the reference category, with a value of *'0'*, whilst the *pre-fed treatment* was assigned a value of *'1'*.

the pattern of *latency* across *probe value*, either when *lever choice* and *contingency* were taken into account, or not. However, the two-way interaction between *pre-feeding treatment* and *contingency* was significant (9162.965 – 9152.180 = 10.785, 1 d.f, p=0.001; Equation 0.102), and an inspection of the coefficients indicated that whilst the responses of both *contingency* groups were slower following *pre-feeding,* compared to *no pre-feeding,* the extent of this difference between the two *treatment* groups was greater for the *4kHz=2pell contingency* group. The addition of the three-way interaction between both of these terms (*pre-feeding* and *contingency*) and *lever choice* also improved the model (9152.180 – 9144.339 = 7.841, 2 d.f., p=0.020; Equation 0.103), indicating that whilst *pre-feeding* slowed the lever press responses of both *contingency* groups, for the *2kHz=2pell* group it slowed the '2-pellet' lever response more than the '1-pellet' lever responses, whilst the opposite was true for the *4kHz=2pell* group: i.e. for this group, *pre-feeding* had a larger effect on the *latency* to press the '1-pellet' lever, than the *latency* to press the '2-pellet' lever.

As in the multilevel analysis of *lever choice*, described above, the *prior treatment* terms were then systematically added to the model. The addition of *unpredictable-housing treatment* group, as a main effect[152], did not improve the fit of the model (9144.339 (Equation 0.103) - 9142.081 (Equation 0.104) = 2.258, 1 d.f, p=0.133); however, when the effect of *unpredictable-housing treatment* group was allowed to vary across *lever choice*, with the addition of the relevant interaction, the fit of the model was substantially improved (9142.081 - 9117.638 = 24.443, 1 d.f., p<0.001; Equation 0.105). An examination of the coefficients indicates that whilst *subjects* previously in the *UHT* group were faster than the *Control* group to record each type of lever response (although this main effect was not significant, of course), the difference between the two *prior treatment* groups was greater for presses made on the '1-pellet' lever. Otherwise, a two-way interaction of *unpredictable-housing treatment* group with *pre-feeding treatment,* and a three-way interaction between these terms and *lever choice,* were explored, but these did not significantly improve the model's fit.

[152] The *Control (predictable housing)* group was assigned the reference category, with a value of *'0'*, whilst the *unpredictable housing treatment (UHT)* group was assigned a value of *'1'.*

The addition of *Enrichment treatment* group (see Chapter 3), as a main effect[153], did not improve the model's fit (9117.638 (Equation 0.105) - 9117.445 (Equation 0.106 = 0.193, 1 d.f, p=0.660), nor did the addition of interactions between this term and both *lever.choice,* and *pre-feeding treatment.*

Finally, as in the multi-level analysis of *lever choice,* above, a few other terms of interest were added to the model. The addition of *pre-feeding counterbalanced group* (i.e. whether *subjects* were first *pre-fed* before their first test session, or before their second), as a main effect[154], explained a significant amount of variance, indicating that those rats who were *pre-fed* before their first test session were overall slower to record a lever press response (9117.638 (Equation 0.105) - 9111.855 (Equation 0.107) = 5.783, 1 d.f., p=0.016). In addition, *Session no.* (i.e. 1 to 6) also exerted a significant main effect, 9111.855 (Equation 0.107) - 9094.042 (Equation 0.108) = 17.813, p<0.001), revealing that lever press responses became slower as *Session no.* progressed. The *mean weight of food pellets eaten in the 1-hour free-feeding test,* when added as a main effect, did not improve the model (9094.042 – 9093.762 = 0.28, 1 d.f., p=0.597). Similarly, the addition of *bodyweight* (as measured the day preceding the start of probe-testing) as a main effect did not explain a significant amount of the variance, and whilst its interaction with *treatment* did (9094.042 – 9072.504 = 21.538, 2 d.f., p<0.001; indicating that heavier rats were faster to record a lever press response when *not pre-fed,* but *latency* to press respond was more similar across *bodyweight* when *pre-fed*), the size of the corresponding coefficient was, as we found in the analysis of *lever choice,* very small (0.001(0.000)) suggesting the effect was of relatively little biological interest.

In the Appendix, Figure 0.52 and Figure 0.55, which plot the residuals from the final fitted model, indicate the assumptions of the model are reasonably well met.

---

[153] The *Non-enriched treatment* group was assigned the reference category, with a value of *'0'*, whilst the *Enriched treatment* group was assigned a value of *'1'*.

[154] *First pre-fed prior to Session 2* was assigned the reference category, with a value of *'0'*, whilst *First pre-fed prior to Session 1* was assigned a value of *'1'*.

Finally, we briefly present the main findings from two subsidiary analyses, each of which only modelled the *latency* data from one type of response, namely presses on the '2-pellet' lever, and presses on the '1-pellet' lever, respectively. For the categorical predictor (x) variables, the choice of reference category remained the same as for the *latency* analysis of all the lever presses, above.

Firstly, the analysis which modelled only those responses on the '2-pellet' lever (n=3,877) again found a highly significant main effect of *pre-feeding treatment* (4350.213 (Equation 0.109) – 4268.844 (Equation 0.110) = 81.369, 1 d.f., p<0.001), with presses on the '2-pellet' lever considerably slower in sessions which followed *pre-feeding,* compared to those which did not. Otherwise, the interactions of *pre-feeding treatment* we explored featuring *contingency,* and *probe value,* did not significantly improve the model, nor did the most relevant main effects and interactions we investigated involving the *prior treatment* groups.[155] Again, the addition of *bodyweight*, as a main effect and in interactions, revealed the same pattern as above, with the (statistical) significance of the *bodyweight*treatment* interaction indicating that the difference in *latency* across *bodyweight* was greater when *not pre-fed*, but with the size of the corresponding coefficient again very small. Otherwise, interactions of *bodyweight* with *probe value,* with or without *treatment,* were not significant. In the Appendix, Figure 0.56 and Figure 0.59, which plot the residuals from the final fitted model, suggest the model's assumptions have been reasonably well met.

Finally, the analysis which modelled only those presses made on the '1-pellet' lever (n=3,688) also found a highly-significant effect of *pre-feeding treatment* (4802.801 (Equation 0.111) – 4722.406 (Equation 0.112) = 80.395, 1 d.f., p<0.001), and a highly-significant two-way interaction between this term and *contingency* (4722.406 (Equation 0.112)– 4708.717 (Equation 0.113) = 13.689, 1 d.f., p<0.001), indicating that whilst the responses recorded on the '1-pellet' lever were slower following pre-feeding for both *contingency* groups, the difference between the two treatments

[155] Actually, *pre-feeding*UHT*probe value* was a significant term (p=0.017), but since the biological significance of this interaction is not especially clear, and it was only one significant term among many tested, it was not included in further models.

was greater for the *4kHz=2pell* group. Otherwise, the interactions of *pre-feeding treatment* we explored featuring *probe value* (with and without *contingency* in the interaction term), did not significantly improve the model, nor did the most relevant main effects and interactions we investigated involving the *prior treatment* groups. Finally, neither *bodyweight,* nor its interaction with *treatment* had a significant effect. In the Appendix, Figure 0.60 and Figure 0.63, which plot the residuals from the final fitted model, suggest the model's assumptions have been reasonably well met.

Reference trial latency

The (log-transformed)[156] mean latency to record a lever press in each type of reference trial (i.e. for those associated with two pellets of food, and for those associated with one pellet of food) were submitted to exploratory analyses (separately, for each trial type), using repeated-measures GLMs, with *accuracy* (i.e. correct, or not) and *treatment* as within-subjects factors, and *contingency* as a between-subjects factor. Each of these analyses found a highly significant effect of *accuracy*, and a significant effect of *treatment* (for the analysis of the '2-pellet' reference tones, this was highly-significant); otherwise, the interaction between *treatment* and *accuracy* neared significance at the 0.05 level when modelling the '2-pellet' reference tones, but not when modelling the latency to respond to the '1-pellet' reference tones, and there were also some significant interactions featuring *contingency* (although we shan't dwell on these in this relatively brief, exploratory analysis) ('2-pellet' reference trials - *treatment:* $F_{1,13}$=15.995, p=0.002; *contingency:* $F_{1,13}$=1.068, p=0.320; *treatment*contingency:* $F_{1,13}$=5.028, p=0.043; *accuracy:* $F_{1,13}$=20.135, p=0.001; *accuracy*contingency:* $F_{1,13}$=0.091, p=0.768; *treatment*accuracy:* $F_{1,13}$=3.777, p=0.074; *treatment*contingency*accuracy:* $F_{1,13}$=0.011, p=0.919; '1-pellet' reference trials - *treatment:* $F_{1,13}$=4.737, p=0.049; *contingency:* $F_{1,13}$=1.590, p=0.229; *treatment*contingency:* $F_{1,13}$=0.319, p=0.582; *accuracy:* $F_{1,13}$=13.999, p=0.002; *accuracy*contingency:* $F_{1,13}$=5.699, p=0.033;

[156] NB Whilst both analyses satisfied formal tests of homogeneity of variance, some of the residuals failed formal tests of normality. This was not readily remedied by any of the transformations we employed, at least not without compromising variance homogeneity (a more important condition to satisfy: e.g. Grafen & Hails, 2002), and so we present the best compromise.

*treatment*accuracy*: $F_{1,13}$=0.988, p=0.338; *treatment*contingency*accuracy*: $F_{1,13}$<0.001, p=0.986). Figure 4.13, which plots these data, summarised by *treatment, trial type,* and *accuracy,* confirmed our inspection of the estimated marginal means from these models: namely that the *pre-fed* group had a longer *latency*, both groups were quicker to record correct responses, and there was a non-significant trend for the *pre-fed* group to have an especially long latency when recording incorrect responses to the '2-pellet' reference tone.

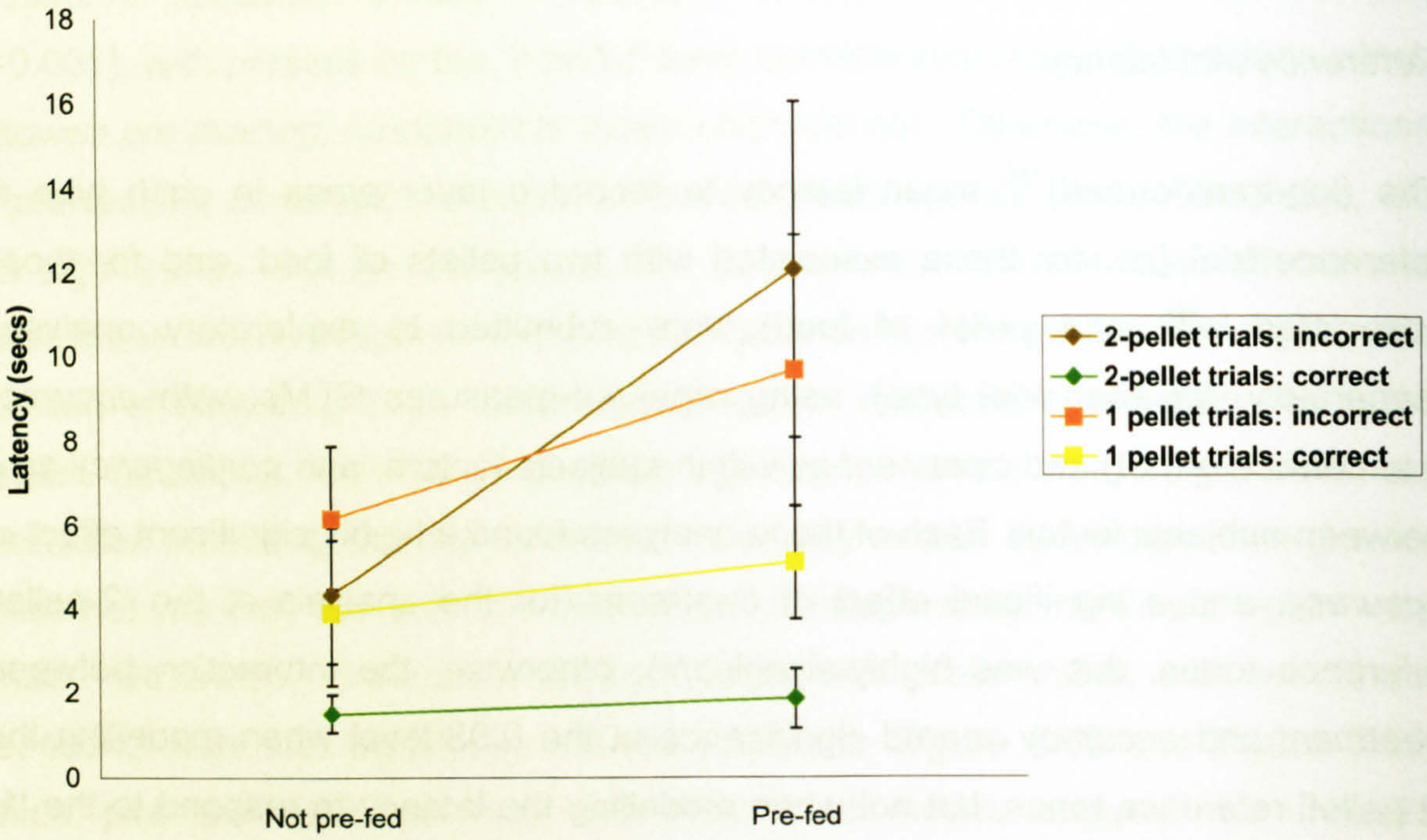

Figure 4.13 The mean latency to press a lever in response to the reference trials (i.e. those in which 2kHz, or 4kHz, were presented, including both reinforced, and non-reinforced trials) in the single-frequency probe sessions, across *treatment,* summarised by trial type (i.e. amount of associated food reinforcement) and also accuracy (i.e. whether the 'correct' lever was pressed in response to that tone, or not) (+/-1SEM).

Free-feeding test

In the free-feeding test, the rats ate a mean average of 9.8 grammes of food pellets (SEM: 0.50; range 5.84g–13.62g). Figure 4.14 plots mean *bodyweight* (as measured each morning of the free-feeding tests) against mean *weight of food eaten* in the free-feeding tests, for each *subject;* these variables were not significantly correlated (*r*=-0.041, d.f.=13, p=0.884). Incidentally, in the operant

probe-testing sessions in the current experiment, the rats averaged, overall, 81.4% accuracy when presented with the reference trials. If they completed all trials in a session (i.e. were not 'timed-out'), then at this level of accuracy, they would receive approximately 3.6g of food pellets.

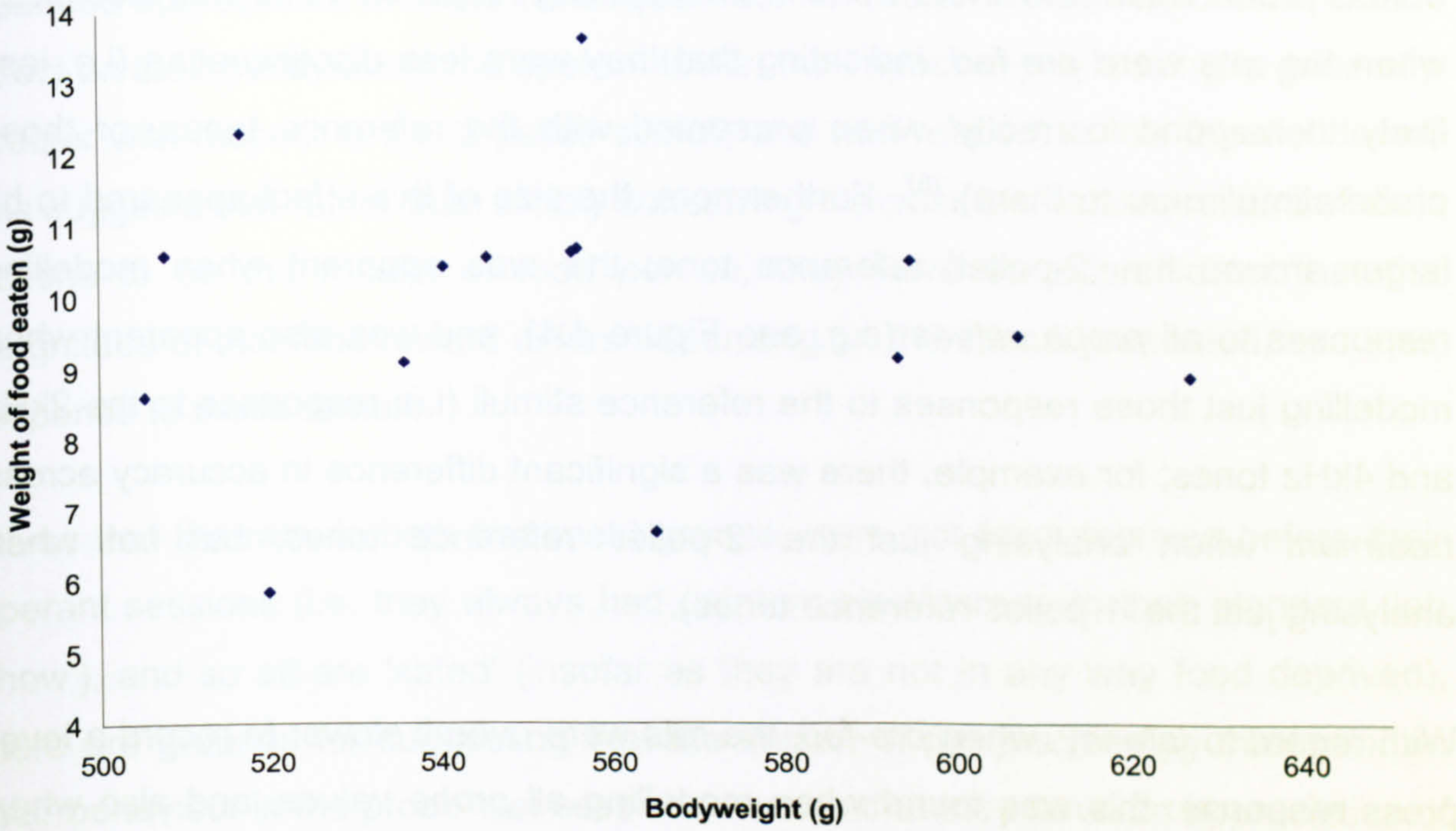

Figure 4.14 The mean weight of food eaten across the three free-feeding test sessions, plotted against mean bodyweight (as measured on the morning of each test session), for each *subject*.

DISCUSSION

In this experiment we varied the level of pre-feeding that rats received prior to two-choice operant discrimination probe tests in which a variety of ambiguous stimuli - variously differing from reference stimuli the rats had previously learnt were reinforced with two different quantities of food - were presented. Before half of these probe-test sessions, the rats were pre-fed with the same type of food they routinely received as reinforcement during their operant sessions (but had rarely received outside the operant chamber), whilst before the remaining sessions, the rats received no such pre-feeding; in both instances, the rats had normal, *ad libitum* access to their regular homecage 'lab chow' prior to the operant tasks.

We found that the *pre-feeding treatment* had a significant effect on various aspects of the rats' behaviour in the probe test sessions. For example, whilst the *pre-feeding* treatment did not have an effect on overall *lever choice* (i.e. when the rats were *pre-fed* they were neither more, nor less, likely to press the lever associated with two pellets of food), it had a significant effect on their pattern of *lever choice* across *probe value:* the linear trend in the resulting response curve was less steep when the rats were *pre-fed*, indicating that they were less discriminating (i.e. less likely to respond 'correctly' when presented with the reference tones, or those probe stimuli near to them).[157] Furthermore, the size of this effect appeared to be larger around the '2-pellet' reference tone; this was apparent when modelling responses to all *probe values* (e.g. see Figure 4.4), and was also apparent when modelling just those responses to the reference stimuli (i.e. responses to the 2kHz and 4kHz tones; for example, there was a significant difference in accuracy across *treatment* when analysing just the '2-pellet' reference tones, but not when analysing just the '1-pellet' reference tones).

With regard to *latency,* when *pre-fed*, the rats were overall slower to record a lever press response: this was found when modelling all *probe values,* and also when modelling only those responses to the reference stimuli. In the former analysis (i.e. of all *probe* values) this difference in *latency* across *treatment* did not interact with *lever choice* (i.e. the *latency* to press the '2-pellet' lever, for example, was not more, or less, affected by the *pre-feeding treatment* than the *latency* to press the '1-pellet' lever), although there was a suggestion of some asymmetry when modelling responses to the reference stimuli alone (the *treatment* slowed responding to '2-pellet' reference tones by a level of greater statistical significance ($p=0.002$) than it slowed responding to '1-pellet' reference tones ($p=0.049$).

Our findings suggest that pre-feeding the rats had a general effect on their performance in the subsequent operant task in keeping with a decline in their 'food motivation'; whilst it's possible that a satiated rat may change its performance in tasks in which food is not used as a reinforcer (e.g. perhaps satiation changes

[157] Somewhat similarly, Echevarria et al (2005) found that rats' accuracy in operant discrimination tasks, using water as the reinforcer, decreased if they were allowed pre-session access to water.

biological functioning in a variety of ways which aren't obviously food-specific), perhaps the most likely explanation of the change in the subjects' behaviour in this particular operant task is that they 'know' what the consequences of their actions are: i.e. they 'know' that if they engage with the task, they may receive food reward. In addition, though, there is a tentative suggestion that they have a nuanced appreciation of these consequences: i.e. rather than an across-the-board effect on their behaviour in the operant tasks, the impact of pre-feeding is generally greater with respect to the '2-pellet' stimuli, and with respect to the '2-pellet' lever; this suggests that rather than simply conceiving the possible consequences of their actions to be the receipt of food *per se*, they have an appreciation that the magnitude of that food reward differs depending on the stimuli presented, and their response to those stimuli.

Given that the rats, in both treatment groups, were not food-deprived before their operant sessions (i.e. they always had uninterrupted access to their standard 'lab chow'), and so all are 'sated' (insofar as they are not in any way food deprived), there are grounds for suggesting that the impact of the *pre-feeding treatment* on their behaviour in the probe-test sessions reveals that they have an appreciation of the *qualitative* difference between the food they receive in the operant chamber and their day-to-day fare; i.e. they have a food-type-specific satiety (e.g. Rolls, 2005) which renders them less engaged in a task rewarded by that *specific type of food.* However, whilst this may be the case, and indeed such an effect has been found elsewhere (e.g. Dickinson et al., 1996), in this instance it's possible (perhaps fairly likely) that when the rats were *pre-fed,* they ate more (of any food type) than is usually the case in the period just prior to the probe test sessions[158]; if so, then they may be *generally* more sated in the *pre-fed* treatment, and therefore might be *generally* less engaged in tasks reinforced by *any* food type.

Our overall findings have two important implications. Firstly, they suggest that the rats employed as subjects in the current experiment – and indeed as subjects in the previous two experiments, using the same, or similar, operant tasks – 'knew'

[158] As mentioned in the Results, they typically ate all the pellets (2g) presented to them in the ten minutes preceding the probe test session.

that their actions in the probe tests could (a) result in the receipt of food, and (b) result in the receipt of quantities of food which differ depending on the stimulus presented, and their response to it. Secondly, they imply that differences in motivational states pertaining to food – i.e. differences in hunger, 'food motivation', etc. – affect performance in operant tasks such as the one employed in this experiment. The veracity of the former conclusion is vital for the integrity of the experimental hypotheses tested in the previous two experiments (i.e. those described in Chapters 2 and 3), whilst the second conclusion alerts us to the possibility that an 'optimistic-style' of responding in tasks such as these *may* reflect a higher level of 'food motivation'.[159] These conclusions are, of course, two sides of the same coin: experimental designs such as these, which employ a two-choice task where one set of contingencies are associated with an outcome 'better' than the other, rely on the fact that the subjects are 'aware' of this difference, but this same 'awareness' renders the experiment vulnerable to changes in how much the subjects value this difference. It may be that a different design, or different analyses, could separate out such differences in utility, but in the current set of experiments, we'd be concerned if changes in 'food motivation' *didn't* change responding.

Whilst the account given in the Results was only descriptive, the charts plotting the change in the rats' accuracy when presented with the reference tones as the probe-test sessions progressed (i.e. Figure 4.6 and Figure 4.7), suggested some interesting patterns. There was a suggestion of two 'phases' as the sessions progressed, each characterised by a bias towards greater accuracy when presented with the '2-pellet' reference tone, and separated at an approximate mid-way point by an attenuation of this net bias. Furthermore, the *not pre-fed* group appeared to be, overall, more accurate in the first (i.e. earlier) 'phase', whilst the accuracy of the two *treatment* groups was more similar in the second (i.e. later) 'phase'. Whilst the following account is speculative, it's nonetheless worth discussing the possible implications of within-session changes in responding, since

[159] It's worth noting that not all indices of operational 'optimism' were affected by pre-feeding: for example, when *pre-fed*, the rats were not *overall* more likely to press either lever; however, there was some asymmetry in responding across *probe value* (with regard to the magnitude of associated food reinforcement), which suggests the potential for such confounds is real.

these may have implications for how we glean information, for experimental designs such as this, in support of hypotheses concerning 'cognitive bias', ambiguity, subjective probability and affect. It's possible that the first phase corresponded to a relatively engaged, food-motivated 'optimism': for example, the *not pre-fed* rats performed to a higher level than the *pre-fed* rats (with regard to overall accuracy) suggesting a role for 'food motivation' in the rats' behaviour, yet both groups still express an ('optimistic') bias towards pressing the '2-pellet' lever. Roughly mid-way through the session, perhaps due to an increase in the rats' level of satiety (since they will have been receiving food reinforcement earlier in the session), and/or due to fatigue, or a general waning of focused attention, this 'optimistic' bias attenuates (earlier for the *pre-fed* rats, speculatively because they start the session closer to their satiation threshold), and is perhaps replaced by a relatively more inattentive 'default' or 'pre-potent' pattern of responding (you may recall we discussed the relationship between the deployment of attentional resources and the ability, or otherwise, to inhibit 'dominant' responses in the Discussion in Chapter 2: typically when attention is less focused on the task at hand, the ability to inhibit 'prepotent' or 'default' responses is compromised); there is little difference in overall accuracy between the two *treatment* groups (unlike in the first 'phase'), yet the '2-pellet' bias is re-established for both groups. This mooted possibility of a temporal change in 'default' responding, occurring earlier when *pre-fed,* gains some credence from the non-significant trend for rats to make incorrect responses when presented with the '2-pellet' reference tone *later* when *pre-fed* (e.g. see Figure 4.13); such long latencies tentatively suggest a higher degree of inattentiveness, and their direction (i.e. towards pressing the '2-pellet' lever) suggests a preponderance of 'dominant' responses in those circumstances. As mentioned above, clearly this is a speculative account, but it at least suggests that it might be worthwhile taking within-session changes into account when gauging the presence, or otherwise, of an 'optimistic' bias.[160]

[160] Incidentally, as an aside, Figure 4.6 and Figure 4.7 indicated an initial improvement in overall accuracy over the first few trials, perhaps as the subjects reacquainted themselves with the task, and made fine adjustments to their responding in the face of initial feedback (i.e. 'got their eye in', to use a metaphor from a different sensory faculty).

We also found that rats in the counterbalanced group *pre-fed* before their *first* probe session (as opposed to *pre-fed* before their second probe session), were overall slower (i.e. across *all* sessions: following pre-feeding or not) to record a lever press response. This somewhat suggests that the apparent devaluation effect evident in the probe tests which immediately followed pre-feeding may have pervaded subsequent test sessions as well, including those not prefaced with such pre-feeding. Such an effect may also be responsible for the overall slowing across *Session number* (i.e. the rats were generally slower to press a lever in the later sessions). In addition, we found that subjects were overall less likely to press the '2-pellet' lever in later sessions (i.e. across *Session number*); this concurs somewhat with the previous two studies, which found that rats were less likely to press the '2-pellet' lever in the later *measurement phases*. It may be that rats are less likely to erroneously bias their responding towards the '2-pellet' lever as they become more discriminating as their experience becomes more extensive; this is a fairly speculative conclusion, although the level of individual variation (as measured by the variance partition coefficient in relatively simple models) has gradually attenuated across the last three experiments, suggesting that the subjects, as a whole, are perhaps becoming more 'expert' as time goes on.

Interestingly, the significant interaction between *contingency*, *pre-feeding treatment* and *lever choice* indicated that the *latency* to press the lever associated with the 4kHz tone differed more across *treatment* than the *latency* to press the lever associated with the 2kHz tone; i.e. whilst being *pre-fed* generally slowed the rats' responses, this was to a greater extent when the lever in question was associated with the 4kHz tone. In Chapter 2 we discussed the psychophysical properties of the probe stimuli we chose for our probe tests (i.e. values ranging from 1.6kHz – 4.4kHz, in 200Hz intervals), noting that more of the *probe values* were likely to sound similar to the 4kHz tone than they were to the 2kHz tone. Therefore, it *may* be that any attentional disengagement from the task, which resulted from being *pre-fed*, would particularly compromise performance with regard to the 4kHz tone, since this reference tone was likely to have been harder to distinguish from *probe values* of similar Hertz frequency; i.e. a 'confident', and quick response when presented with the 4kHz stimulus may require greater attentional resources than a similarly 'confident' response when presented with the 2kHz tone.

Finally, as in the study described in the last chapter, some behavioural differences were found across *prior treatment* grouping, in a direction in keeping with those uncovered in earlier experiments. For example, rats previously in the *enriched* treatment (i.e. those rats who received additional 'enrichments' in the experiment outlined in Chapter 3) were significantly more likely to press the '2-pellet' lever in the current study, and indeed there was an overall non-significant trend in the same direction in the previous ('environmental enrichment') experiment, with the difference between the two treatment groups increasing towards the end of that study. It's conceivable that the differences observed in the present study reflect a persistent effect of the *enrichment* treatment, but it's also possible that it reflects individual differences which persist somewhat regardless of treatment (i.e. there was a trend for the *enriched* rats to be *generally* more likely to press the '2-pellet' lever in the previous study, both before, and during, the *enrichment* treatment). We also found some differences in performance between the rats who were previously in the *unpredictable-housing treatment (UHT)* in a prior study (i.e. that described in Chapter 2), and those who were previously in the *Control* group in that experiment. The differences in operant performance between those groups seemed to have steadily attenuated since that particular treatment ended: in the previous study (i.e. that described in Chapter 3) the *UHT prior treatment* group were significantly more likely to press the '2-pellet' lever, and were overall quicker to record a lever press response, as indeed they had been towards the end of the study in which the *unpredictable housing treatment* was originally applied, however the difference in *lever choice* across these *prior treatment* groups (i.e. the *UHT* and *Control* groups) attenuated over the course of the 'enrichment' study, whilst in the current experiment, there were no significant differences in *lever choice* at all; similarly, whilst the rats who had been in the *UHT* group were overall faster to record a lever press response in Chapter 3, this effect was less comprehensive in the current experiment.

Here, we end the experimental chapters concerning the two-choice operant task we have developed with rodents; we will provide a general overview of this work in our final chapter, but next turn to a different paradigm, with a different study species: namely a test of affect-related foraging bias in domestic chicks.

CHAPTER 5

EMOTIONAL STATE AND FORAGING BEHAVIOUR IN CHICKS (GALLUS GALLUS DOMESTICUS)

INTRODUCTION

Aposematism and multimodal signalling

Animals which are toxic, or unpalatable, to potential predators sometimes have visual properties which render them conspicuous against the backgrounds on which they are commonly encountered: for example, the red colouration of a leaf-dwelling ladybird (family Coccinellidae), the striped patterning of certain wasps (family Vespidae), and so on (e.g. Cott, 1940; Edmunds, 1974). When it is thought these cues function in averting predator attacks, they are termed *aposematic*[161]. As well as these visual signals, when under threat of attack many such animals employ cues in other sensory modalities, constructing an aposematic signal which is *multimodal*[162]: for example, they may buzz, rattle, emit characteristic smells or tastes, and so on (e.g. Rowe & Guilford, 1999a; Rowe & Skelhorn, 2005).

Aposematism has been the subject of considerable theorising and experimentation by biologists seeking to understand its evolution (e.g. Joron, 2003; Ruxton et al., 2004); more specifically, multimodal aposematism has attracted interest from those wishing to account for why such an elaborate signal, which is presumably more costly to produce, should have evolved in preference to a simpler signal in just one sensory modality (e.g. Rowe & Guilford, 1999a). To this end, a number of theories, not necessarily mutually-exclusive, have been proposed to account for how the various components of a complex aposematic signal contribute tactically to

[161] derived from the Greek for "away" and "sign" (Joron, 2003)

[162] i.e. involving more than one sensory modality (also sometimes called *multisensory*); the term *multicomponent* is used when a signal consists of more than one part, or feature, but not necessarily more than one sensory modality, and all of these terms (*multi-modal, -sensory & -component*) are generally considered instances of *complex signaling* (e.g. Hebets & Papaj, 2005; Partan & Marler, 2005; Rowe, 1999).

realising the presumed strategic aim of averting predator attacks; some of these are summarised in Table 5.1.

The elicitation of unlearnt foraging biases

Rowe and her colleagues have conducted a number of experiments which provide support for the last of the hypotheses listed in Table 5.1, namely that the non-visual component of an aposematic signal may elicit, in a potential predator, an unlearnt bias against attacking the visual component (e.g. Jetz et al., 2001; Rowe & Guilford, 1996; Rowe & Guilford, 1999a, 1999b; Rowe & Skelhorn, 2005). They have used domestic chicks (*Gallus gallus domesticus*) as subjects, and studied their preference between food of a colour commonly found in aposematic insects, such as red or yellow, and food of a colour not commonly associated with aposematism, such as green. These experiments have found that when chicks are presented with certain non-visual stimuli, such as pyrazine smells, buzzing noises and quinine tastes, their foraging behaviour is biased against attacking red or yellow food items (and therefore towards green food items, for example); these biases are either attenuated, or not apparent at all, in the absence of the non-visual cues. In addition, the novelty of the non-visual stimulus seems to be important in eliciting this change in behaviour: for example, presenting one of a number of different odours (including some not associated with aposematism) biases chicks' foraging behaviour away from attacking yellow food, but only if the odour is novel; if it is not, there is no bias (Jetz et al., 2001). Furthermore, as well as against colours commonly found in aposematic insects (e.g. red or yellow), biases are also elicited against novel colours (Jetz et al., 2001; Marples & Roper, 1996; Rowe & Guilford, 1999a), and against colours which are conspicuous against their background (Lindstrom et al., 2001).

Improve detection speed
A complex signal may speed detection of the animal from its background, and certainly there is evidence that a compound stimulus is detected more quickly than one which is simpler.

Circumvention of blocked sensory channels
Signalling over more than one sensory channel improves the chances of detection if a channel is blocked, for example by other signals in the environment or ambient stimuli.

Improve discrimination accuracy
A complex signal improves the likelihood the animal will be accurately discriminated from other possible stimulus identities.

Predator-specificity
The various components of a signal may be aimed at different types of predator.

Increase of information content
A signal with more than one component can provide more information than a signal consisting of just one.

Within-group summation
When two stimuli are learnt separately, an animal's response to their subsequent presentation is greater if they appear together.

Between-groups summation
A compound stimulus is learnt more quickly than a stimulus consisting of one of the components alone.

Potentiation
Presentation of a non-visual and visual component together may enhance the learning of the latter, so that when it is subsequently presented alone, a greater response is elicited than if it was learnt in isolation.

Increase hesitancy to attack
Some components may increase the predator's hesitancy to attack, increasing the likelihood the predator will more closely inspect the animal and accurately assess its toxicity, and/or move on to attack other stimuli.

Decrease habituation
A predator may habituate more slowly to a multicomponent stimulus compared to one consisting of just one component.

Elicitation of unlearnt bias
The presentation of a non-visual component may elicit an unlearnt bias against attacking stimuli with certain visual qualities.

Table 5.1 Summary of some of the theories proposed as to how the various components of a complex aposematic signal contribute to averting attacks by potential predators (for reviews and discussion, see Partan & Marler, 2005; Rowe, 1999; Rowe & Guilford, 1999a).

Rowe & Guilford (1999a) proposed that these non-visual stimuli may induce a greater hesitancy to attack, and as a consequence, potential food items may be more carefully-inspected. In nature, where there may be genuine toxicity or unpalatability, such inspection may result in a more accurate judgement of an item's value (i.e. there may be a speed-accuracy trade-off); alternatively, inducing hesitancy may simply result in the animal moving on to attack other stimuli more quickly. In the experimental situation, where the nutritional status of foods of different colours is often the same (e.g. Jetz et al., 2001; Lindstrom et al., 2001; Marples & Roper, 1996; Rowe & Guilford, 1999b; Rowe & Skelhorn, 2005), judgements of their profitability may be more greatly swayed by their visual properties, which receive greater scrutiny: if some are associated, *a priori*, with a relatively negative value, then they may be more likely to be rejected. Rowe & Guilford (1999a) proposed an additional mechanism that may operate alongside one of increased hesitancy: namely, that as a given visual cue (e.g. yellow) is encountered in a new context (e.g. with a buzzing noise), then the significance, or value, of that visual cue (or the *Gestalt* of which it is a part) might be perceived in isolation from past experience with that cue. This 'perceptual isolation hypothesis' again assumes some innate asymmetry with regard to the significance of different visual properties: some, when perceived as part of a new percept, may be judged *a priori* as having a more negative value, and/or may be more likely to be 'perceptually-isolated', i.e. they may be more integrated into a compound whole, which is perceived as a new stimulus, and it is this novelty itself which is aversive (Bronson, 1968; Marples & Kelly, 1999).

Affective state & cognitive bias

Our previous discussions of the links between affective state and certain cognitive processes (e.g. in Chapter 1, page 15 onwards) suggests a mechanism via which presentation of the non-visual stimuli used in these experiments may result in the observed change in the chicks' foraging behaviour, one which is not exclusive of the hypotheses outlined above. Namely, the non-visual stimuli may induce a negative change in the chicks' affective state, and a biasing of some of the cognitive processes involved in foraging behaviour, resulting in a reduced tendency to attack stimuli with certain visual qualities.

Why might the non-visual stimuli induce such a change in affective state? Whilst it is true that some of the treatments used in these experiments, like quinine (Rowe & Skelhorn, 2005), are likely to be intrinsically aversive to chicks, it seems this is not the case for all of the stimuli used in these experiments, such as pyrazine odour (Rowe & Guilford, 1999a, p.657). However, as mentioned earlier, novelty of these stimuli is an important determinant of their effect on foraging bias (Jetz et al., 2001), and chicks are typically fearful of novel stimuli (e.g. Jones, 1996), at least after imprinting has occurred (Rogers, 1995, p.91). Therefore, on account of their novelty, these stimuli may induce a negative, perhaps anxiety-like change in the chicks' affective state.

As discussed previously, negative affective states are associated with the characteristic biasing of certain cognitive processes in humans: for example, a greater tendency to judge ambiguous events, or stimuli, as having a negative outcome, or significance (for a review, see, for example, Paul et al., 2005). A typical paradigm used to investigate colour-based foraging biases in chicks involves presenting subjects with a simultaneous choice between a number (usually two) colours of food, both of which are novel (e.g. Rowe & Skelhorn, 2005). Each colour of food is therefore ambiguous in the sense that it has not been encountered before (although the extent of its perceived novelty might depend on its visual similarity to food the chick has encountered before: i.e. the extent of any generalisation), and the outcome of an attack on either is therefore unknown, in that it is uninformed by experience. However, the food which shares visual qualities (e.g. red) with some toxic prey items might, through an unlearnt, *a priori* conceptualisation on the part of the chick, have a different range of *possible* outcomes, or values. When the chick is in a negative affective state, the tendency to select a value from this range which is negative might be greater. In the case of food items with visual qualities not associated with aposematism (such as green), the range of anticipated outcomes, *a priori*, may be narrower, and thus the worse anticipated possible outcome might have a more neutral valency; hence, a foraging bias against attacking food with visual qualities shared with some aposematic insects would result.

In addition to these changes in judgements, in humans, a negative change in affective state is often accompanied by greater attention being directed to possible sources of threat in the environment (as discussed in Chapter 1, and in the following reviews: C. MacLeod, 1999; Paul et al., 2005). If the same is true of other species, such as chickens, then we might expect them to direct greater attention towards foods which, *a priori*, have certain visual qualities more likely to have a threatening significance (such as being toxic). The predictions from such a change in attentional resources are likely to be similar to Rowe & Guilford's (1999a) hypothesis regarding hesitancy (described earlier)[163]: attacks on certain foods would be less ballistic, and more careful assessment of these foods may result in a greater tendency to categorise them as unprofitable, and/or to move on more quickly to the next crumb.

We investigated the role of affect on chicks' foraging behaviour by employing a treatment specifically designed to induce a negative affective state in chicks (rather than a laboratory-based analogue of a non-visual component of an aposematic signal). We chose a three minute period of social isolation, which has been validated as producing an anxiety-like state in chicks, inducing distress vocalising which is attenuated by the administration of anxiolytics (Feltenstein et al., 2004; Warnick et al., 2006). We predicted that chicks receiving this treatment would subsequently show a greater bias away from attacking red food, and towards attacking green food, compared to a control group receiving no such isolation.

METHOD

Subjects & housing

32 domestic chicks (*Gallus gallus domesticus*), all hatched on the same day, were assigned to either an experimental group (n=27; 14 male, 13 female) or a group of

---

[163] Although, in an experimental situation in which a novel non-visual stimulus (e.g. buzzing noise) is delivered into the ambient environment, according to the hypothesis proposed by Rowe & Guilford (1999a) any hesitancy induced by that stimulus may result in more careful inspection of *all* foods (this wouldn't be the prediction in a natural situation, where the various components of a signal all come from the same source, namely the aposemtatic animal); in the negative affect hypothesis proposed here, attention may be targeted more asymmetrically, towards foods with (*a priori*) more threatening significance.

'buddy chicks' (see below; n=5; 3 male, 2 female)[164]. All chicks were housed in two cages measuring 100(L) x 50(W) x 50(H) cm, with experimental and buddy chicks in separate cages. Food (brown chick starter crumbs) and water were available in these cages *ad libitum.* Heat lamps and room heaters maintained a cage temperature of around 24-25°C, and the lighting cycle was set at 15 hours on / 9 hours off, using fluorescent lights with no UV component. All experimental procedures were conducted in the room in which the chicks were housed, during the light phase. Experimental chicks were weighed each day of the experiment, and all chicks were checked for health daily. All chicks were re-homed on small free-range holdings at the end of the experiment.

Test arena

The test arena (see Figure 5.1) was housed in a cage of the same dimensions as the homecage. At one end, 20cm along its length, a rectangular section was separated from the rest of the cage by a wire mesh spanning its width. This section housed two buddy chicks during all habituation and test sessions: they were changed every three trials, and had food and water *ad libitum.* The remaining, larger section constituted the test arena, and had a white floor divided by fine pencil lines into 80 rectangles of equal area; these were used when distributing the crumbs prior to each test session (see below). In addition, the pencil line divisions spanning the width of the test arena were used to later record the chicks' movements from video footage taken during the test session: these marked out ten floor areas of 8cm width, with Floor Area 1 closest to the buddy chicks, and Floor Area 10 furthest.

[164] Chicks who were subjectively-assessed as being the least 'viable' were assigned to the 'buddy chick' group, as these were thought most likely to fail to habituate to foraging alone in the test arena.

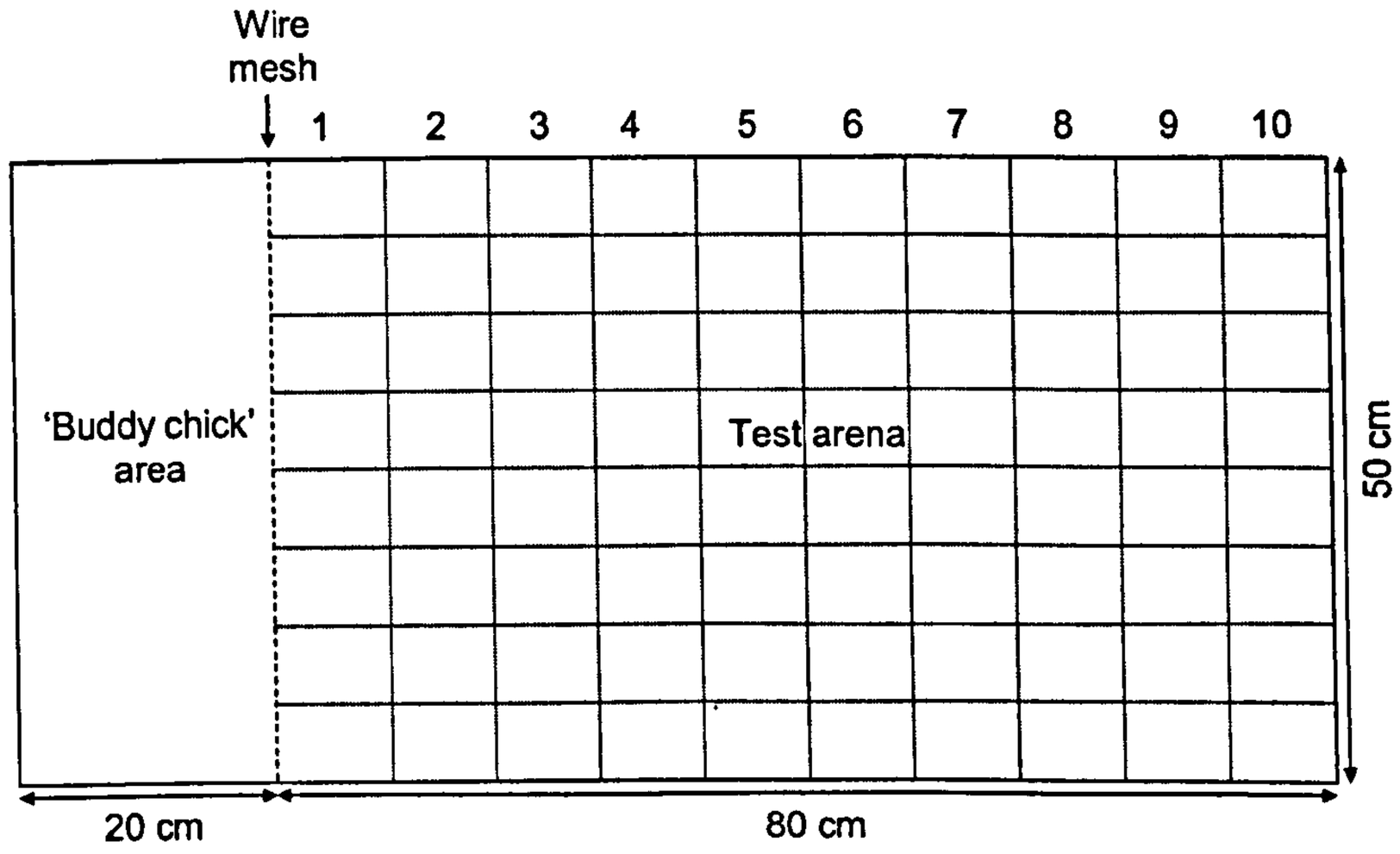

Figure 5.1 Floor plan of the test arena. The numbers (1-10) towards the top of the diagram denote the ten Floor Areas, spanning the width of the test arena, used to record the chicks' movements during the test session.

Habituation

On Day 1 and Day 2 (taking hatch day as 'Day 0') the experimental chicks received a series of sessions designed to gradually habituate them to foraging alone in the test arena. In these sessions, brown chick starter crumbs were scattered across the floor of the arena.

On Day 1, the chicks had three sessions in which they were placed in the test arena in groups of three for 5 minutes, and a final session in which they were placed there in pairs for 4-5 minutes.

On Day 2, the chicks had one session in which they were placed in the test arena in pairs, then three sessions in which they were placed there individually. These sessions lasted 5 minutes, except for the final individual session, which lasted 4 minutes. Prior to each session the chicks were food deprived for *c.*90 minutes. By their final session, all chicks were eating readily in the test arena.

Test crumb preparation

After being sieved to a consistent size, 150g of brown chick starter crumbs were sprayed evenly with either 2ml of Supercook red food dye or 0.5ml of Sugaflair spruce-green food dye, each diluted to 90ml with tap water. They were left to dry and then sieved once more to ensure size consistency.

Test sessions

Each experimental chick received one test session, on Day 3. Before the test session, 20 red crumbs and 20 green crumbs were placed on the test arena floor in a pseudorandomised pattern. This pattern was subject to the following limitations: of the 80 squares on the floor, each quadrant of squares (i.e. comprising 20) had 5 red and 5 green crumbs. Within that quadrant, distribution of these 10 crumbs was random, but with a maximum of one crumb per square. Within a treatment group, each chick had a different pseudorandomised pattern of crumb distribution.

Prior to each test session the chicks were food deprived for *c.*90 minutes. A test session commenced with the chick being placed in the centre of the test arena (in Floor Area 6, facing the buddy chicks – see Figure 5.1). For the first 16 crumb 'attacks' (any contact between beak and crumb constituted an attack), the colour of the crumb was recorded, together with the outcome of that attack (i.e. whether the attack ended in the crumb being eaten or not). Sessions were terminated when either 16 crumbs had been attacked, or 15 minutes had passed, whichever came first. Recording was undertaken manually by an observer, and subsequently cross-checked against video footage from each session, ensuring as accurate a record as possible.

Treatments & order of testing

Experimental chicks were assigned[165] to either a 0-minute isolation group ("0-min" group; n=13; 7 males, 6 females) or a 3-minute isolation group ("3-min" group; n=14; 7 males, 7 females). All chicks were housed, with conspecifics, in a food deprivation cage for *c.*90 minutes prior to their test session. Chicks in the 0-min group were taken straight from this food deprivation cage to the test arena for their test session. Chicks in the 3-min group, however, were taken from the food deprivation cage and placed alone in an isolation cage for three minutes immediately prior to being placed in the test arena for their test session. The food deprivation cage and isolation cage were of the same design as the homecages, but simply lacked food. The food deprivation cage and isolation cage were approximately the same distance from the room's light source above, and were both heated by a heat lamp of the same wattage (150W).

To counterbalance the effects of any vocalising from the isolated chicks on subjects in subsequent test sessions, the test session of half the chicks in each treatment group followed a test session with a chick of their own treatment group, with the other half following a test session with a chick in the other treatment group.

Chick vocalisations

An audio recorder (Sony ICD-B300) was placed in the roof of the isolation cage to record any vocalisations made by the isolated chick; the number of calls made by the chick was later taken from this recording.

Video analysis

The chicks' movements around the test arena were recorded from video footage taken during their test sessions, using 'The Observer 5.0' software (Noldus, 2006).

---

[165] Chicks were assigned to the two groups in a manner which kept the following as balanced as possible: distribution of sexes; the order in which they were selected for initial weighing and marking (in case this order was sensitive to differences in behavioural traits, size, etc.); the ink colours used to mark them.

Specifically, their location within the ten Floor Areas (see Figure 5.1) of the grid drawn onto the floor of the test arena was recorded from the beginning of each test session (the moment they were placed into the test arena, in Floor Area 6) to its end (when they had attacked 16 crumbs); they were judged to have entered a given Floor Area when both legs stood in it.

Data analysis

Statistical analyses were conducted using SPSS 14.0 (SPSS Inc., 2006). Data were submitted to parametric analyses only when the assumptions of those tests were satisfied, with transformations applied as appropriate. When conducting repeated-measures ANOVAs with a within-subjects factor which had more than 2 levels (i.e. $k>2$), as in the last chapter, we follow the advice of Quinn & Keough (2002), and reject the null hypothesis if either the adjusted univariate output, or the multivariate output, reports significance at the 0.05 level.

RESULTS

26 of the 27 experimental chicks attacked 16 crumbs within the 15 minute criterion. The chick who failed to do so was a female in the 0-min group, and data from this subject was excluded from all subsequent analyses. The resulting sample size was therefore: 0-min group n=12 (7 male, 5 female); 3-min group n=14 (7 male, 7 female). Sex had no effect on any of the measures, and was therefore omitted from the analyses.

The proportion of crumbs attacked, by colour

As Figure 5.2 shows, the 0-min group attacked more green than red crumbs, whilst the 3-min group attacked more red than green, and this difference between the two treatment groups was significant: the 0-min group attacked a significantly smaller proportion of red crumbs than the 3-min group ($t_{24}=2.087$, $p=0.048$). Figure 5.3, which further subdivides this data into the first and last eight attacks, indicates that both treatment groups attacked proportionately fewer red crumbs, and proportionately more green crumbs, in the last eight attacks compared to their first

eight attacks. For each crumb colour, the proportion of available crumbs attacked over the first eight was compared to the proportion of available[166] crumbs attacked over the last eight attacks: this found a significantly greater proportion of green crumbs were attacked over the last eight compared to the first eight attacks (W=-2.599, *N*=26, *p*=0.009), but there was no significant difference with regard to red crumbs attacked (W=-1.398, *N*=26, *p*=0.162).

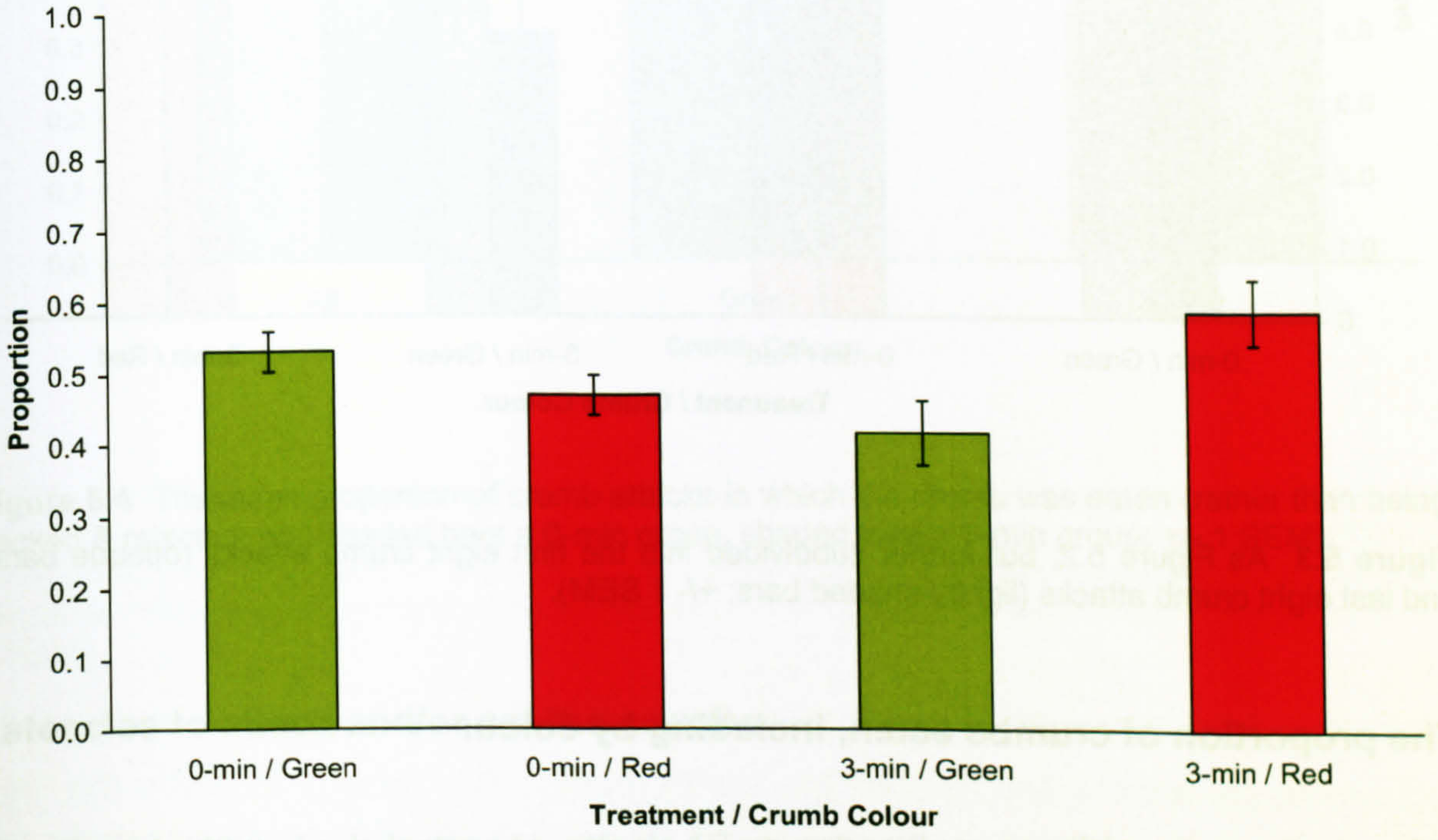

Figure 5.2 The mean proportion (by treatment group & crumb colour) of crumbs attacked by chicks (data from 16 crumb attacks, which was the criterion for test session termination; +/- 1 SEM).

[166] i.e. uneaten in the first eight attacks. This index isn't ideal: whilst it is a measure of those *physically* available, it doesn't reflect any change in the biological significance of those crumbs pecked and rejected (i.e. these crumbs *may* be less likely to be attacked again, perhaps on account of the chicks' movements around the arena, away from areas they have previously visited). However, defining the number of crumbs available for attack over the last eight as those unattacked (regardless of whether eaten or not) in the first eight yielded similar results (for green crumbs, W=-2.707, *N*=26, *p*=0.007; and for red crumbs: W=-1.057, *N*=26, *p*=0.291).

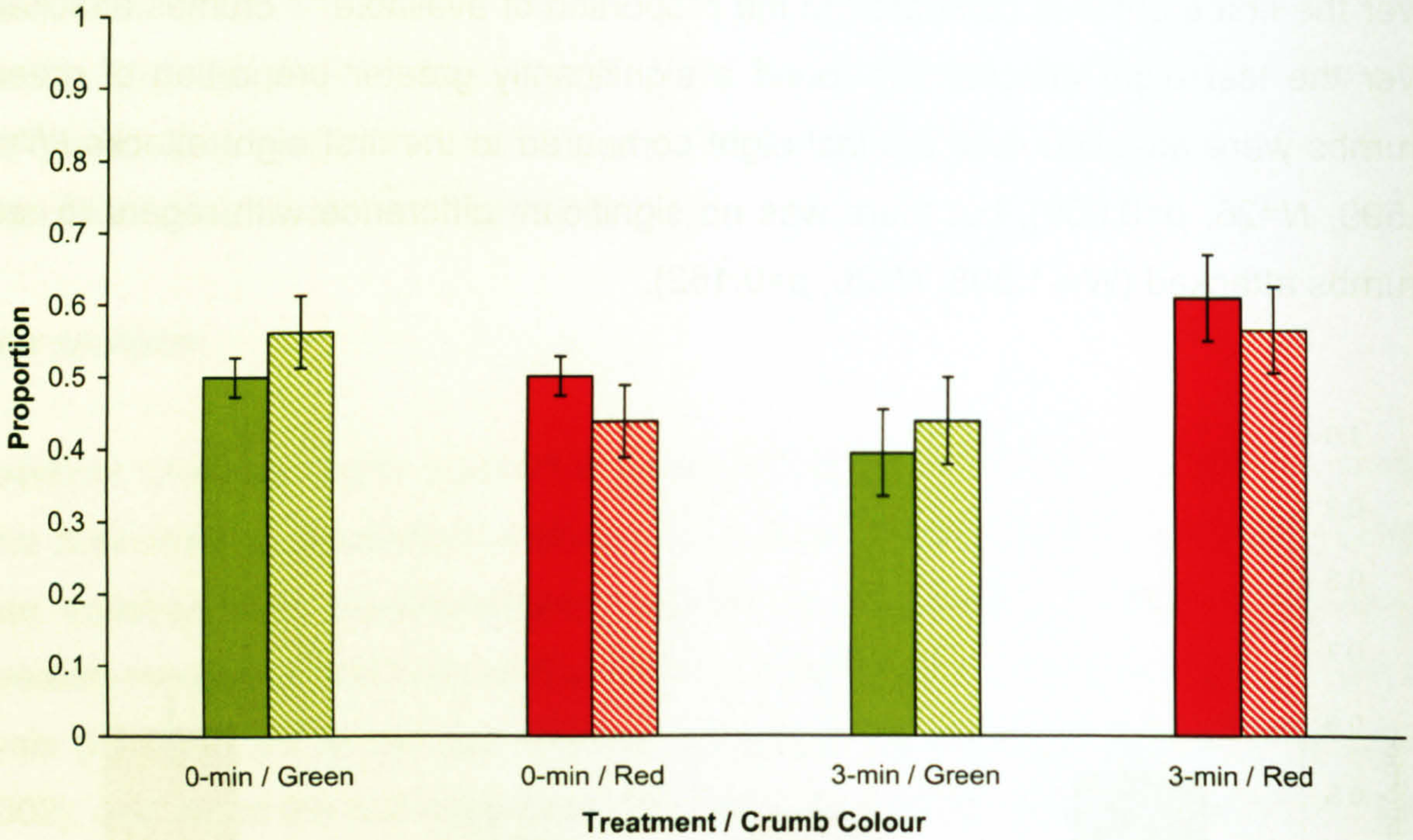

Figure 5.3 As Figure 5.2, but further subdivided into the first eight crumb attacks (opaque bars) and last eight crumb attacks (lightly-shaded bars; +/- 1 SEM).

The proportion of crumbs eaten, including by colour

All chicks ate at least one crumb (the fewest eaten was 6, the most 16), but not all chicks ate crumbs of both colours (one chick in the 3-min group ate no green crumbs; another chick in the same treatment group ate no red crumbs). Whilst Figure 5.4 shows that the 0-min group ate a greater proportion of the crumbs they attacked (regardless of colour) than the 3-min group, this difference was not significant (using arcsine-transformed data: t_{24}=1.408, p=0.172).

A further analysis by colour (see Figure 5.4; each analysis excluded data from any chick who did not attack a crumb of that colour) found no significant differences between or within groups (proportion of green crumbs eaten once attacked, by treatment group: W=143.5, N_1=12, N_2=13, p=0.126; proportion of red crumbs eaten once attacked, by treatment group: W=163, N_1=12, N_2=13, p=0.729; by colour, 0-min group only: W=133, N_1=12, N_2=12, p=0.259; by colour, 3-min group only: W=172, N_1=13, N_2=13, p=0.852).

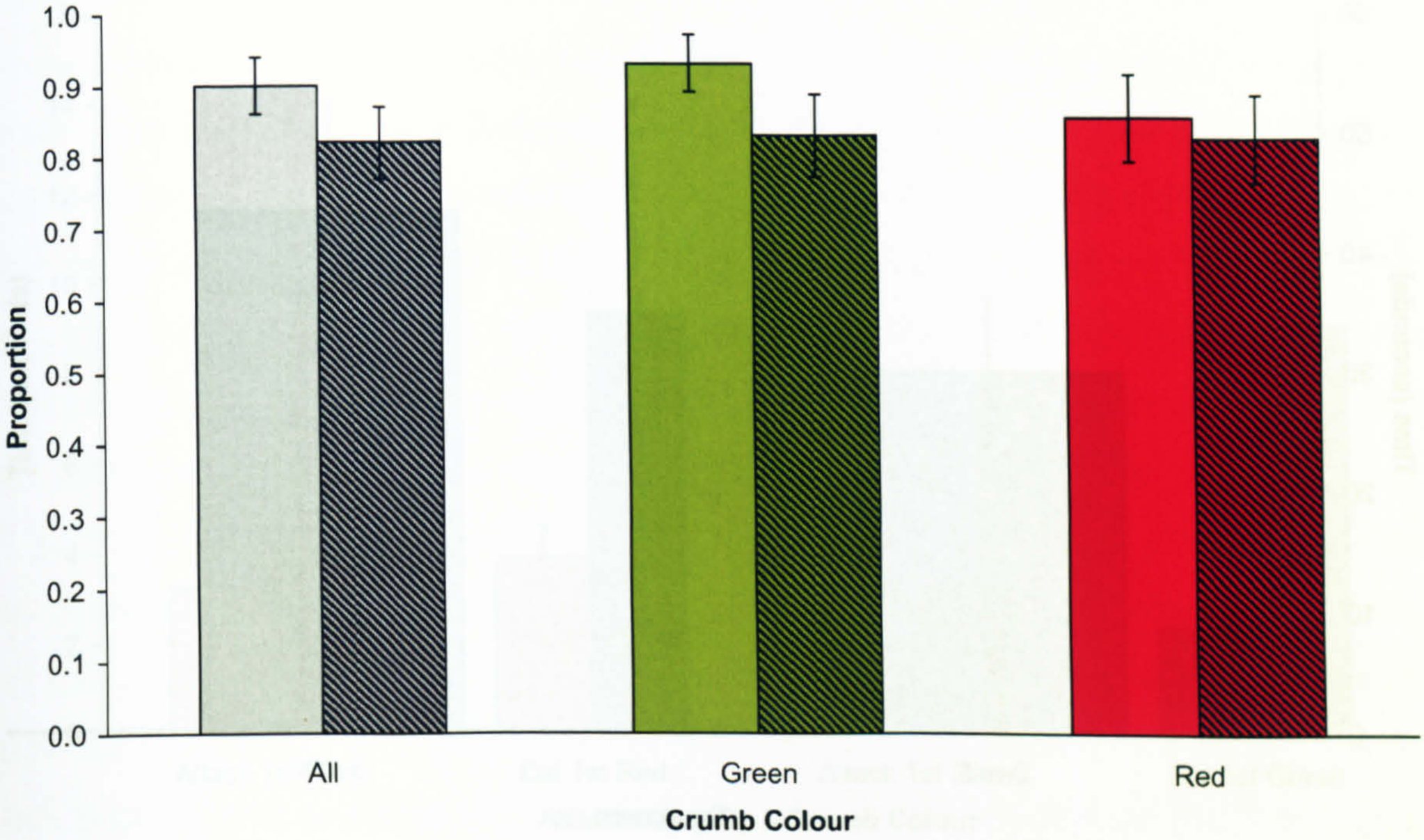

Figure 5.4 The mean proportion of crumb attacks in which the crumb was eaten (rather than being pecked & rejected; non-shaded bars = 0-min group, shaded bars = 3-min group; +/- 1 SEM).

Latencies to attack and/or eat the crumbs

The 3-min group took longer to attack 16 crumbs (thus reaching the criterion for termination of the test session) than the 0-min group (see Figure 5.5), but this difference was not significant (using log-transformed data: $t_{24}=1.214$, $p=0.237$).

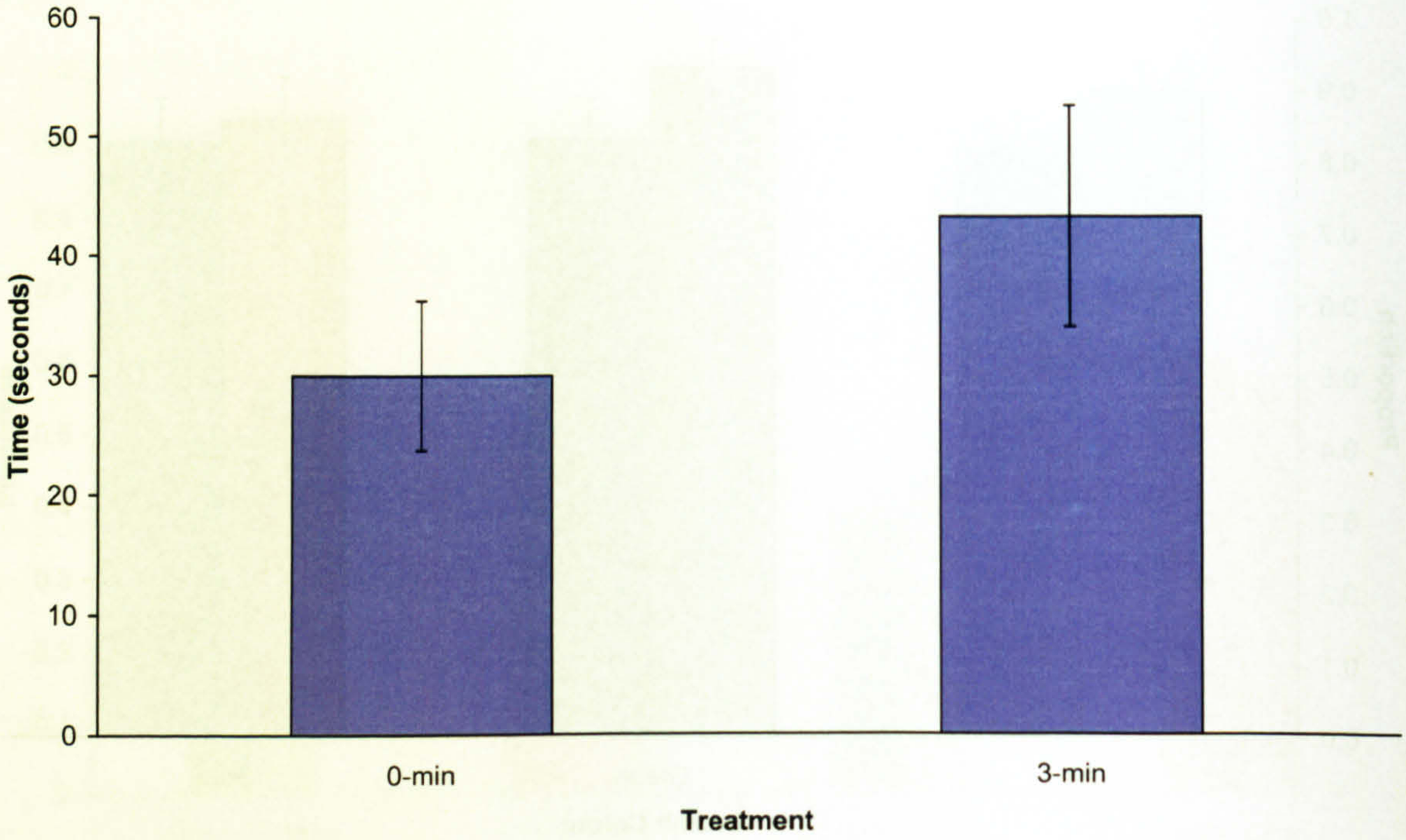

Figure 5.5 The chicks' mean latency to attack 16 crumbs (the criterion for test session termination; +/- 1 SEM).

As Figure 5.6 shows, the 3-min group also took longer to attack and/or eat their first crumb (of both colours), than the 0-min group. The latency to make the first attack (regardless of whether that crumb was eaten) was analysed using a repeated-measures GLM, with crumb colour as a within-subjects factor, and treatment as a between-subjects factor. After first log-transforming the data, no significant main effect of either colour or treatment was found ($F_{1,23}$=2.041, p=0.167; $F_{1,23}$=2.086, p=0.162, respectively), nor a significant interaction between the two factors ($F_{1,23}$=0.032, p=0.860). In addition, the latency to eat the first crumb was analysed in the same way. This found a significant main effect of treatment ($F_{1,22}$=5.171, p=0.033; i.e. the 3-min group took a significantly longer time from the start of the test session to first eat crumbs of either colour), but there was no significant main effect of colour ($F_{1,22}$=0.780, p=0.387), nor any significant interaction between the two factors ($F_{1,22}$=2.379, p=0.137).

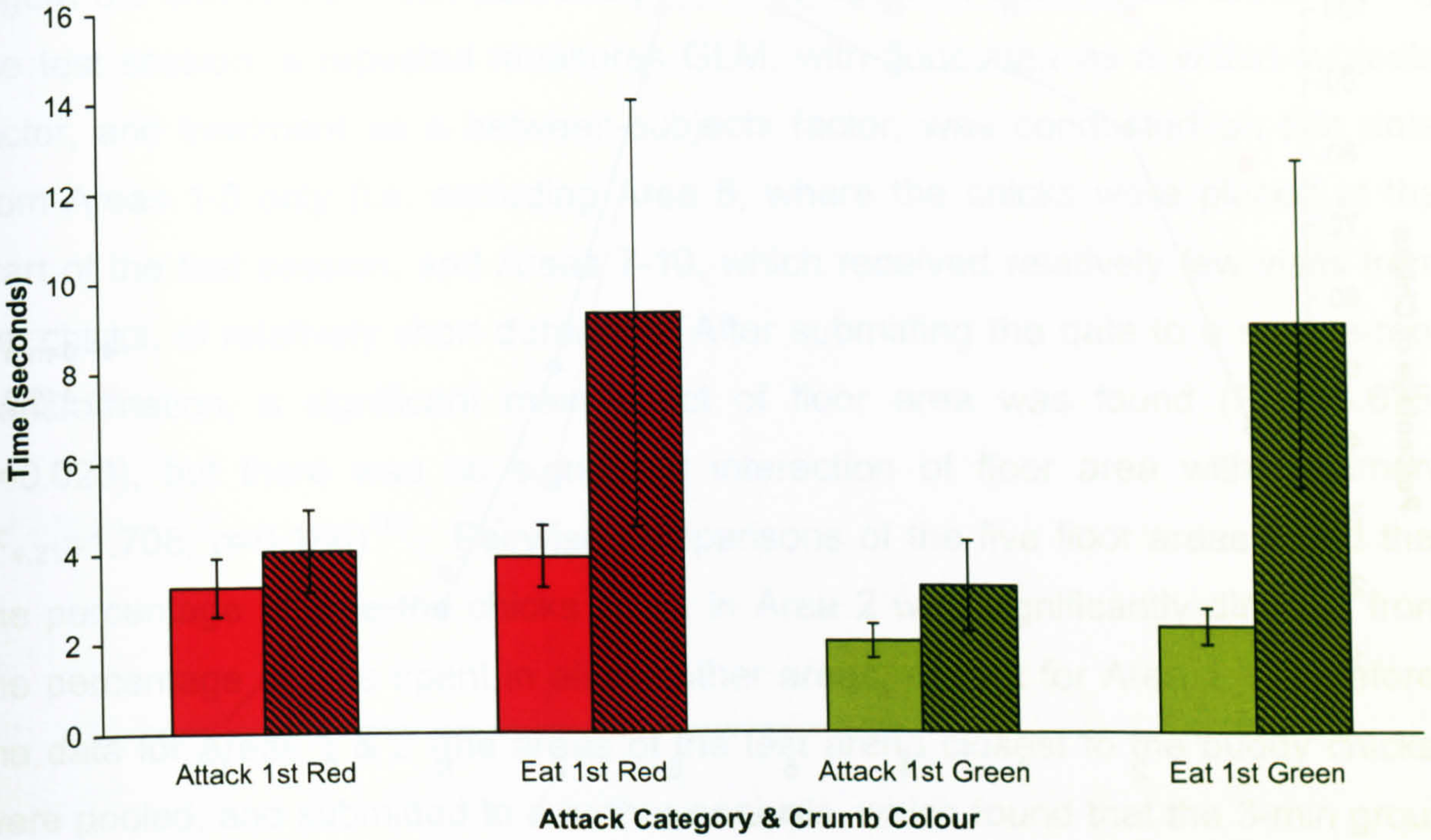

Figure 5.6 The chicks' mean latency to attack / eat the first crumb, by treatment group and crumb colour (NB an 'attack' is any contact between beak and crumb, regardless of whether that crumb is subsequently eaten, or pecked & rejected; for each bar, only data from chicks attacking / eating a crumb of that colour are included; non-shaded bars = 0-min group, shaded-bars = 3-min group; +/- 1 SEM).

Proximity to buddy chicks in the test arena

Analysis of the video footage taken during the test session showed that not all chicks entered each of the ten areas of the test arena floor, with a greater percentage of chicks in the 3-min group entering the areas closest to the buddy chicks (see Figure 5.7). Figure 5.8 indicates that the chicks in the 3-min group also had a slightly shorter mean latency to enter the areas closest to the buddy chicks, and a higher mean latency to enter the areas furthest away (see Figure 0.67 in the Appendix for an alternative method of representing this data).

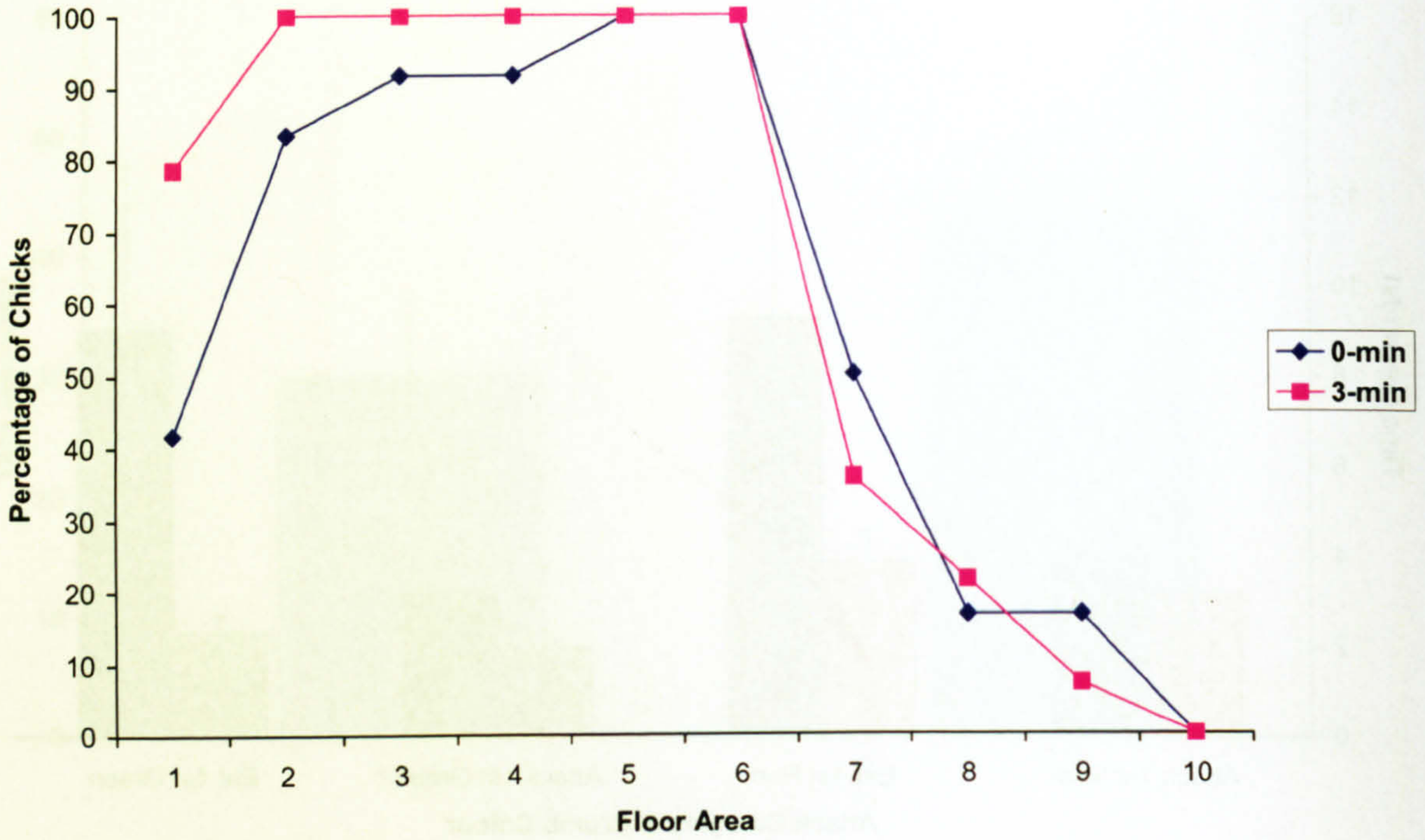

Figure 5.7 The percentage of chicks in each treatment group entering the various floor areas of the test arena during their test session (see Figure 5.1 for floor plan of test arena; N.B. the chicks were placed in Floor Area 6 at the start of the test session, facing the buddy chicks).

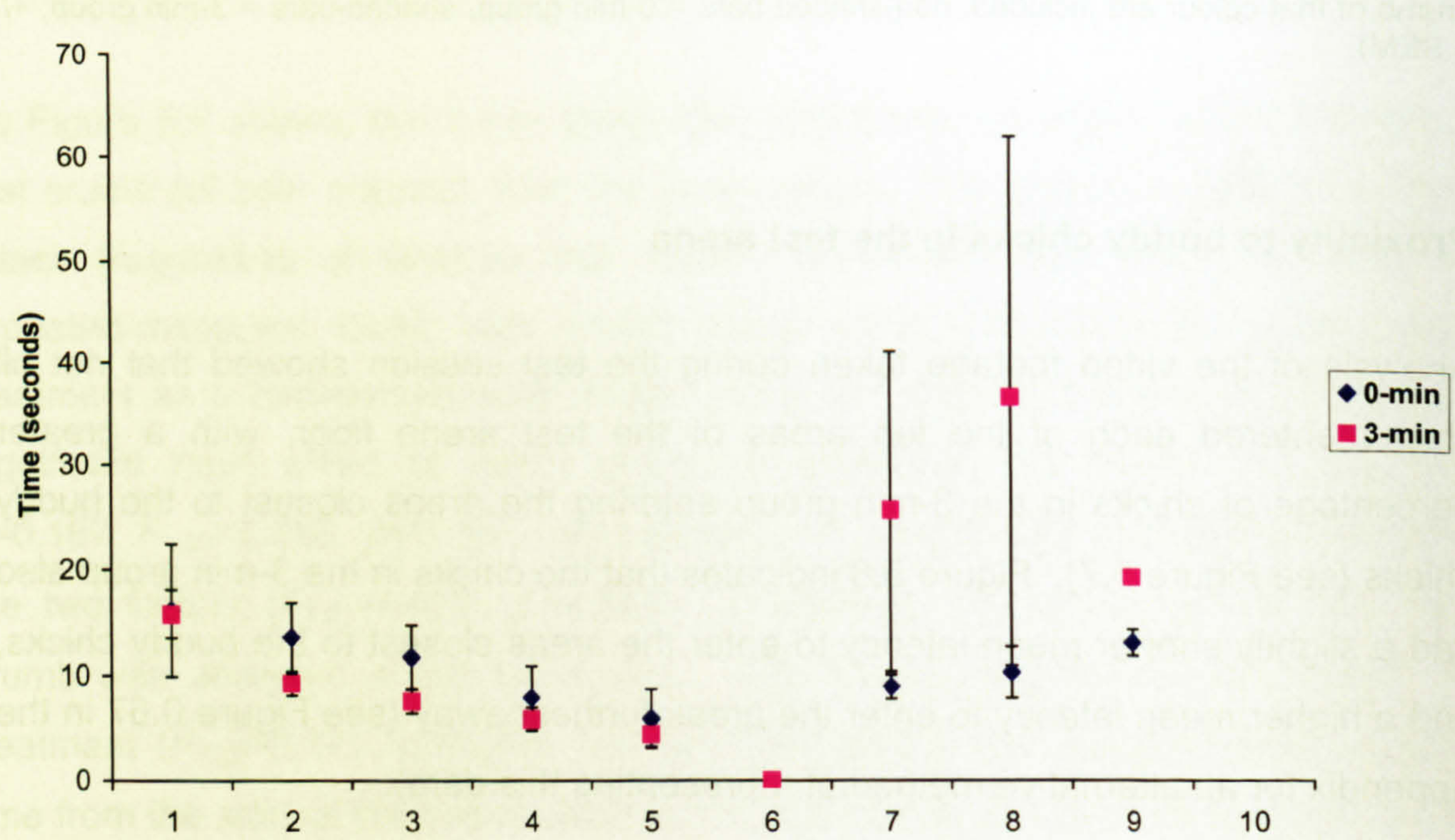

Figure 5.8 The chicks' mean latency to enter each floor area of the test arena, by treatment group (see Figure 5.1 for floor plan of test arena; N.B. the chicks were placed in Floor Area 6 at the start of the test session, facing the buddy chicks; data is taken only from chicks entering a particular floor area; +/- 1 SEM).

Figure 5.9 shows the mean percentage of time spent in each of the areas during the test session: a repeated measures GLM, with floor area as a within-subjects factor, and treatment as a between-subjects factor, was conducted on this data from Areas 1-5 only (i.e. excluding Area 6, where the chicks were placed at the start of the test session, and Areas 7-10, which received relatively few visits from the chicks, of relatively short duration). After submitting the data to a square-root transformation, a significant main effect of floor area was found ($F_{4,21}$=3.675, p=0.020), but there was no significant interaction of floor area with treatment ($F_{4,21}$=1.708, p=0.186)[167]. Pairwise comparisons of the five floor areas found that the percentage of time the chicks spent in Area 2 was significantly different from the percentage of time spent in all the other areas, except for Area 1. Therefore, the data for Areas 1 & 2 (the areas of the test arena closest to the buddy chicks) were pooled, and submitted to a further analysis, which found that the 3-min group spent a significantly greater percentage of their test session time in Areas 1 & 2 than the 0-min group (t_{24}= -2.271, p=0.032).

The number of calls made during social isolation

The mean number of calls made by the chicks in the 3-min group during the 3 minute period of social isolation was 340.36 (min: 153, max: 428; S.E.M. = 18.63). There were no significant correlations between the number of calls made and any of the following behavioural variables in the subsequent test session (all latency data was first log-transformed): the proportion of red crumbs attacked (r =0.320, d.f.=12, p=0.265); the latency to attack 16 crumbs (r = -0.352, d.f.=12, p=0.217), the latency to attack the first red or green crumb (red: r = -0.035, d.f.=12, p=0.905; green: r = -0.011, d.f.=11, p=0.972;), the latency to eat the first red or green crumb (red: r =0.092, d.f.=11, p=0.766; green: r =0.215, d.f.=11, p=0.480;); the percentage of test session time spent near the buddy chicks (i.e. in Areas 1 & 2; r = -0.459, d.f.=12, p=0.099). See Figure 0.71 - Figure 0.75 in the Appendix for plots of these data.

[167] In both instances, we quote the multivariate output.

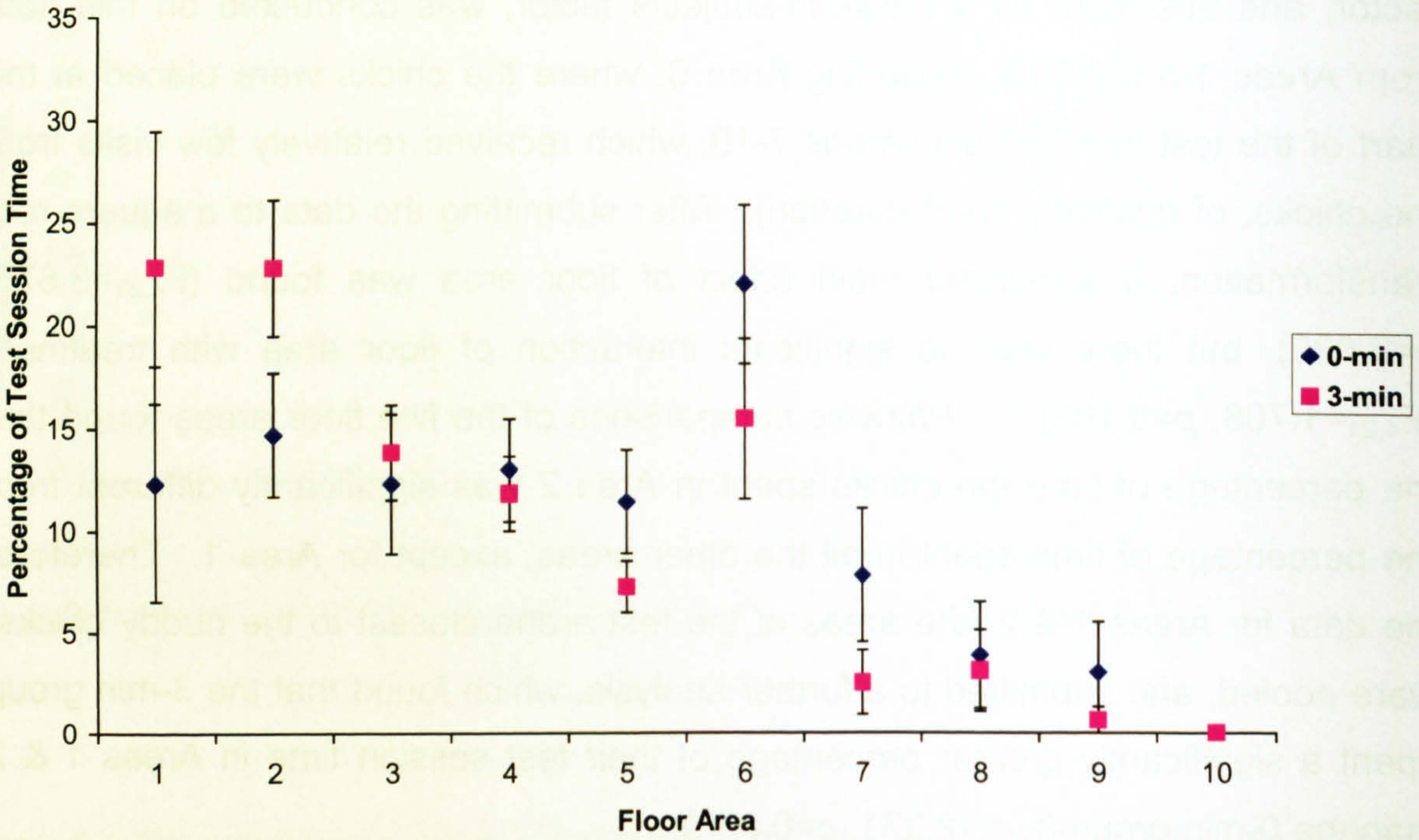

Figure 5.9 The mean percentage of test session time the chicks spent in each floor area of the test arena, by treatment group (see Figure 5.1 for floor plan of test arena; N.B. the chicks were placed in Floor Area 6 at the start of the test session, facing the buddy chicks; test session duration differed between chicks, and was determined by their latency to attack 16 crumbs; +/- 1 SEM).

Additional data

See the Appendix for the following additional charts:

- The mean proportion of attacked crumbs which were eaten (rather than pecked and rejected) over the first eight (Figure 0.64) and last eight (Figure 0.65) attacks.
- By sex: the percentage of chicks entering the various Floor Areas of the test arena (Figure 0.66), their mean latency to first enter these areas (Figure 0.68 & Figure 0.69), and the mean percentage of test session time they spent in these areas (Figure 0.70).

- The mean order, within the 16 crumb attacks made, chicks attacked / ate the first crumb (Figure 0.76), and the mean crumb (Figure 0.77), by crumb colour and treatment group; and the mean percentage of red and green crumbs attacked in each of the ten Floor Areas by chicks in the 0-min group (Figure 0.78), and by chicks in the 3-min group (Figure 0.79).

DISCUSSION

We tested the hypothesis that a treatment designed to induce a negative change in the emotional (affective) state of chicks, namely a three minute period of social isolation, would bias their subsequent foraging behaviour away from attacking red food - a colour commonly found in aposematic insects - and towards attacking green food. Surprisingly, our results indicated the opposite: chicks socially-isolated for three minutes were significantly more likely to attack red food crumbs in a subsequent foraging test than a control group receiving no such social isolation.

Clearly, our hypothesis has not been supported, but the results are no less intriguing for it, and below, we consider the following *a posteriori* hypotheses: an unanticipated change in the emotional valence of the chicks in the 3-min group; a difference in arousal, distraction and cognitive loading between the two groups; and a difference in their level of neophobia.

An unanticipated change in emotional valence

Our original hypothesis assumed that any negative change in affective state induced by the three minute period of social isolation would persist into the subsequent foraging test session, where its effect on foraging behaviour (via a biasing of the cognitive processes involved) would be manifest. However, it is possible that any such negative shift may, on entry to the test arena, be attenuated to a more neutral valence, or changed further into one which is positive. The chicks in the 3-min group are taken from an environment which is putatively aversive, being novel and socially-isolated (Feltenstein et al., 2002), to one which

is familiar (they have been gradually habituated to readily forage alone in the arena), relatively safe (they have had no experience of being harmed there), and allows close visual contact to be re-established with conspecifics (the two buddy chicks behind the wire mesh) and also the experimenter (to whom they may have imprinted over the preceding three days since hatching). It is possible that this positive change in circumstances produces, via a process of appraisal (e.g. Scherer, 1999), a positive change in the chicks' affective state, and this successive contrast may render the affective state of the chicks in the 3-min group as more positively-valenced in the test arena than the chicks in the 0-min group (see Rolls (2005; e.g. pages 11-15) for a discussion of 'relief-like' states following the termination of a punishing[168] stimulus).

The possible consequences of this are as follows: domestic chickens are descended from the omnivorous Red Jungle Fowl (*Gallus gallus*), whose diet includes fruits, berries, seeds, nuts, as well as a variety of insects and other invertebrates (Klasing, 2005). A generalist feeder such as this will encounter a diverse range of potential food items, with a variety of sensory properties and nutritional values. Some of these items will have visual properties rendering them conspicuous against the backgrounds where they are commonly encountered, for example being coloured red (e.g. Schaefer et al., 2006; Schmidt et al., 2004). It is possible they may have evolved such conspicuity because there is an adaptive advantage in being noticed or accurately discriminated by potential predators: i.e. it forms at least part of a signal conveying information about nutritional value (e.g. Guilford & Dawkins, 1991; Willson & Whelan, 1990). The nutritional value of red items is likely to be either particularly good, such as certain fruits and berries, or especially bad, such as certain toxic or unpalatable insects (Batesian mimics aside) (e.g. Cott, 1940; Gamberale-Stille et al., 2007). Since many aposematic animals are coloured red, we hypothesised foraging chicks in a negative affective state would behave in a manner consistent with them judging red items as more likely to belong to the category of toxic or unpalatable prey than green items.

[168] Rolls defines a punishing stimulus as "anything that an animal will work to escape or avoid, or that will suppress actions on which it is contingent". In domestic chicks, a novel, socially-isolated environment is likely to fit such criteria (e.g. Feltenstein et al., 2002; Jones, 1996; Suarez & Gallup, 1983).

However, the opposite may also be true: chicks in a more positive affective state may tend to judge such items as belonging to the category of very profitable food, i.e. they might show a bias in the opposite direction (Isen, 1999). Indeed, there is some evidence that chicks make colour-based foraging choices on the basis of whether the stimuli presented are insects or not. Gamberale-Stille & Tullberg (2001) found that when domestic chicks (*Gallus gallus domesticus*) were presented with insects painted either red or green, they preferred to attack the green stimuli. However, when presented with artificial 'fruits' (actually small balls of pastry, *c.*4mm diameter) painted with the same colours, the chicks showed no such preference.

Whilst the proportional data from red and green crumb attacks supports the hypothesis presented above, it receives less support elsewhere: compared to the 0-min group, the chicks in the 3-min group do not eat a greater proportion of the red crumbs they attack, and it does not explain the greater latency of the chicks in the 3-min group to eat the first crumb: this hypothesis would predict a similar, or even shorter latency, particularly with respect to red crumbs.

Arousal, distraction and cognitive loading

Alternatively, or in addition, the treatment used in this experiment may have altered the cognitive loading of mechanisms which would otherwise be involved in foraging decisions. All chicks come to the test session having had a period of food deprivation (*c.*90 minutes), and this, together with their previous habituation to foraging in the test arena, may predispose them to forage readily on entry to the test arena (novelty of crumb colour aside). The chicks in the 3-min group, however, have also had a period of social deprivation, during which the need to re-establish close social contact may have become increasingly salient. Assuming, for a moment, that these premises are correct, we might expect to see a difference in the behavioural priorities of the two experimental groups of chicks in the test arena: feeding would be a particularly high priority for the chicks in the 0-min group, whilst for the chicks in the 3-min group, there is a concurrent, heightened need to establish and maintain close social company as quickly as possible. Indeed the data supports this: the chicks in the 0-min group are significantly

quicker to eat their first crumb, and there is a trend for them to complete their session (i.e. attack 16 crumbs) more quickly, and to eat more of these crumbs, compared to the chicks in the 3-min group; the latter group, in turn, spend a significantly greater percentage of their test session time staying close to the buddy chicks.

This may have implications for foraging behaviour for the following reasons. Visual attention is often characterised as consisting of 'bottom-up' and 'top-down' processes: the former is drawn towards salient aspects of a visual array (e.g. on account of hue, brightness, shape, movement, and so on), but can instead be directed to task-specific stimuli by the latter, more 'executive' cognitive input (e.g. Connor et al., 2004). In humans, there is evidence that when this top-down input is compromised, for example under conditions of greater cognitive load, or in patients with certain brain lesions, visual attention is more easily swayed by salient and otherwise irrelevant distracters (D'Esposito & Postle, 2000; Lavie et al., 2004). More generally, affective factors, such as high emotional arousal, stress, or pain can also impair performance in cognitive-based tasks in a variety of species, distracting attention, and increasing the likelihood of 'default' responses being performed (e.g. Mendl, 1999; Sneddon et al., 2003); indeed we have discussed such phenomena in previous chapters. For the chicks in the 3-min group, some of the cognitive apparatus which would otherwise be involved in directing visual attention in a task-specific manner may instead be occupied with establishing and maintaining close contact with the buddy chicks and/or be distracted by a generally higher level of arousal. If the red crumbs used in this experiment are more visually salient than the green crumbs[169], then their attention may be more greatly captured by these stimuli, which are therefore more likely to become the target of attacks. The chicks in the 0-min group, on the other hand, are less preoccupied with

169 Without further empirical work (e.g. objective assessment of the visual properties of the coloured crumbs and their background in the test arena (via spectrometry, for example), cross-referenced with the known biology of the chick's visual system; e.g. Bennett et al., 1994; Cuthill et al., 1999), we can say little about the relative salience of the different colours of crumb used in this study, other than noting that objects of this size (c. 2mm diameter) are more likely to be identified on the basis of achromatic cues, such as brightness; chromatic cues, such as hue and saturation, tend to make a larger contribution to the discrimination of larger objects (e.g. Osorio et al., 1999). As a general point, the literature relating to chicks' colour-related foraging *preferences* is rather contrary and inconsistent, possibly due to variation in the design of stimuli (including their size, hue, saturation, brightness, etc.), the lighting conditions and other contextual details, subjects' past visual experience, genetic variation, and so on (e.g. Ham & Osorio, 2007; Rowe & Guilford, 1999a).

gaining close social contact, and therefore have greater cognitive resources involved in directing their attention. Under these circumstances, a tendency to attack conspicuous stimuli may be tempered by a circumspect caution against doing so (e.g. Lindstrom et al., 2001). Interestingly, the pattern of crumb attacks across the test session provides some support for such an interpretation: across both treatment groups, chicks attacked a significantly higher proportion of green crumbs during the last eight attacks compared to the first eight. A pattern such as this would be expected if all chicks take some time to 'settle-down', i.e. to progressively focus more cognitive resources on the foraging task at hand following a period of adjustment to the test arena environment (John Skelhorn, personal communication; see also Regolin et al. (1995) for an example of a cognitive task in which chicks' performance was affected by social isolation and the opportunity for reinstatement).

Stimulus generalisation and neophobia

Our final hypothesis concerns the extent to which the novel green and red crumbs differ, visually, from the brown crumbs previously encountered (Domhnall Jennings, personal communication). If the red crumbs are perceived as being more similar to the brown crumbs, perhaps on account of hue, brightness, and/or the relative saturation of the dyed colour with respect to the original brown of the crumb, then the likelihood of a chick attacking a green crumb will depend, in part, on their level of neophobia (Marples & Kelly, 1999). Therefore, if the three minute period of social isolation renders chicks, at least for a time, more neophobic (Bronson, 1968), then their behaviour may reflect the extent to which the sensory properties of the novel crumbs are generalised to those encountered before (Jetz et al., 2001; Marples & Roper, 1996; Rowe & Guilford, 1999a); if the chicks perceive the red crumbs as being more similar to brown crumbs than green crumbs are, they would be more likely to avoid attacking green crumbs.

Again, the proportional data from red and green crumb attacks offers support for this hypothesis, but elsewhere support is lacking: we would predict that for the chicks in the 3-min group, their latency to attack or eat the first red crumb would be lower than their latency to attack or eat the first green crumb, and that the

proportion of red crumbs they attack which are eaten would be greater than the proportion of green crumbs attacked which are eaten, but we found no such differences. Perhaps more importantly, other studies which have employed treatments which, *a priori*, would be expected to induce a degree of neophobia (such as novel smells and aversive tastes) have found a preference *against* attacking red crumbs (e.g. Rowe & Guilford, 1996; Rowe & Skelhorn, 2005). More generally, objective quantification of the visual properties of the crumbs (i.e. via spectrometry), cross-referenced with the known biology of the chicks' visual system, would provide very useful information to assess the validity of such hypotheses (e.g. Cuthill et al., 1999; Ham & Osorio, 2007; Osorio et al., 1999).

Comparison of results with an earlier pilot study

Before we conclude our discussion, and suggest experiments designed to address some of the methodological and interpretational issues raised here, it's important to note that an earlier pilot study, conducted in the same lab but adopting a slightly different design, yielded results which offer some support for our original hypothesis. In that study, chicks which had been socially-isolated for one minute attacked significantly fewer red crumbs in a subsequent foraging test session, than a control group receiving no such isolation. As well as being socially-isolated for just one minute, as opposed to the three minutes used in the present study[170], the social isolation took place in a covered, dark box, i.e. these chicks entered the test arena having come from a considerably darker environment than the chicks in the control group. In contrast, in the present study chicks were socially-isolated in a cage of the same physical design as the food deprivation cage (where all chicks were housed for *c.*90 minutes prior to their test session), in very similar light and temperature conditions[171]; these cages had a wide, finely-wire-meshed front allowing ambient light to enter.

---

[170] We used a three minute period of social isolation as we thought it important to use a treatment which had been well-characterised and validated as a method of inducing an anxiety-like state in domestic chicks (e.g. Feltenstein et al., 2004; Warnick et al., 2006)

[171] Of course, these environmental conditions will not have been precisely the same, and an argument can be made for moving the chicks in the 0-min group into the 'social isolation cage' for three minutes prior to their test session as well, but this time with conspecifics present. However, it is not only the three minutes of social isolation which induces an anxiety-like,

It is possible that these methodological differences between the present study and the pilot might underlie the differences in their results. For example, as the visual system of the socially-isolated chicks in the pilot study re-adapts to the lighter ambient conditions of the test arena, this may have consequences for how the two crumb colours are perceived. For instance, the dark-adapted human eye is sensitive to lower levels of light, so that when one emerges from a dark environment into a lighter one, the scene seems very bright and 'washed out'[172]; our sensitivity to contrast improves as we stay in that environment, and our visual system becomes increasingly light-adapted (Snowden et al., 2006). In addition, the peak sensitivity of the dark-adapted eye is shifted towards the blue end of the wavelength spectrum, so that maximum sensitivity is around the area of the spectrum we perceive as green, in contrast to the light-adapted eye, which has a peak sensitivity around yellow (this phenomenon is called the 'Purkinje effect', or 'Purkinje shift'; Gregory, 1998). So, whilst the evidence cited relates to human vision (reflecting the bias of the published literature), it illustrates that there is the potential, at least, for the difference in foraging behaviour of the two treatment groups to reflect differences in the way the red and green crumbs are perceived, perhaps due to differences in the contrast of the two crumb colours against their background, or a change in the peak wavelength sensitivity of the visual system of chicks recently placed in a dark environment.

Alternatively, or in addition, it may be the differing duration of social isolation which is responsible for the differences in behaviour: in the pilot, the socially-isolated chicks' preference for attacking green crumbs might be due to the mechanism we originally hypothesised, i.e. a negative change in affective state, and a biasing of some of the cognitive processes involved in foraging behaviour, so that greater attention is focused on possible sources of threat in the environment, and/or there is a greater tendency to judge the consequences of attacking them as negative. If the chicks' cognitive processes are biased as described, then for their effect on

---

or 'stressed', state in chicks, but the novelty of the environment in which they're socially-isolated as well (Feltenstein et al., 2002). Therefore, care might have to be taken not to expose the chicks in the 0-min group to such novelty (e.g. through prior familiarisation), lest the difference in treatment effect between the two groups be attenuated.

[172] Although the extent of this effect depends on how long one has spent in the dark: i.e. one's level of dark-adaptation.

behaviour to be manifest, they would need to be involved in the foraging decisions: directing attention appropriately, and so on. As the period of social isolation lengthens, however, the need to re-establish and maintain close social contact with conspecifics may become an increasing high behavioural priority, such that when these chicks are given the opportunity to do so, more of their cognitive resources are recruited to this end; i.e. the cognitive-biasing first hypothesised may still have occurred, but it exerts no effect on foraging behaviour.

Conclusion and future experiments

Whilst the three *a posteriori* hypotheses we have presented here are not mutually-exclusive, it is the second hypothesis, regarding arousal, distraction and cognitive loading, which seems to fit the data better. As our above discussion of an earlier pilot study illustrates, however, the lack of support for our original hypothesis in the present experiment does not necessarily mean that the predicted changes in the chicks' affective state and biasing of cognitive processes did not occur: it may be that our experiment was simply insensitive to them.

We conclude by suggesting some experimental designs which might resolve some of the issues discussed above. Firstly, whatever treatment is employed to induce a change in affective state, it is important any such change persists into the test session. One approach would be to use a treatment designed to induce a longer-lasting change in affective state, perhaps akin to a pervasive mood, which persists as the chick moves between housing and test environments (this is the strategy we adopted in Chapters 2 and 3, when studying operant responding in rats). For example, the homecage environment and husbandry regime of the chicks could be manipulated in a manner designed to induce a more positive, or negative, affective state. Some care would have to be taken as to the choice of such manipulations, though: varying social density in the homecage (e.g. Marx et al., 2001), for example, could result in some chicks more actively seeking the company of conspecifics in the test arena, distracting their attention from foraging (see above discussion); likewise, the provision of dust-bathing material (e.g. Sanotra et al., 1995; Weeks & Nicol, 2006), or different objects which can be pecked and explored (e.g. Jones, 1982; Jones et al., 2004; Nicol, 1992), risks varying too

greatly the visual experience of the respective treatment groups, leading to problems of interpretation in such a visually-based test design (Miklosi et al., 2002). Some manipulations may be less controversial, though: such as provision of perches (e.g. Brake et al., 1994), or adjustments to the level of cover (e.g. Leone et al., 2007; Newberry & Shackleton, 1997).

Alternatively, the treatment could be a manipulation of the test session environment itself, inducing a change in affective state specific to that session. Given the gregarious nature of domestic chicks (e.g. Feltenstein et al., 2002; Jones & Williams, 1992), one obvious type of manipulation would be to adjust the social environment, perhaps by changing the number of buddy chicks. For example, one could employ a 2 X 2 design in which chicks are habituated to forage in the test arena with either buddy chicks present, or absent. Then, in the test session, this regime would continue for half the chicks (equally represented in each of the two habituation groups), whilst for the other half the presence of buddy chicks would be changed from that encountered during habituation (so that they are either present in the test session when they had not been previously, or *vice versa*). Assuming, for a moment, that the presence of conspecifics induces a positive change in affective state (e.g. Feltenstein et al., 2002), the hypothesis would thus be that chicks foraging in the company of buddy chicks during the test session would attack more red crumbs than chicks foraging in their absence[173]; the 2 X 2 design would control for the effect of novelty *per se* (i.e. control for the effect of *any* change in the number of buddy chicks). As with all designs in which the treatment occurs during the test session, though, there is a risk of the chicks' attention being drawn away from foraging and towards whatever aspect of the environment has been manipulated by the experimenter to try and change the chicks' affective state: be it the presence of conspecifics, sounds, novel objects, and so on. It would therefore be important to calibrate the intensity of the treatment, so that it successfully changes the chicks' affective state, but does not excessively distract

[173] In fact, one might predict that chicks who had been habituated to forage in the test arena in the presence of buddy chicks, but did not during their test session, might show the strongest avoidance of red (if their affective state during the test session correlated with the negative contrast of this change in circumstances); similarly, those chicks who had been habituated to forage in the test arena in the absence of buddy chicks, but who then underwent their test session in their presence might be predicted to exhibit the strongest preference for attacking red. The colour preferences of the other two treatment groups might lie somewhere in between these two groups.

the attention of the chick away from foraging; in practice, though, this might be a challenging objective to meet.

Similarly, if one were to again employ the type of treatment used in the present experiment, namely a period of social isolation just prior to the test session, then by adjusting the length of this treatment, it may be calibrated so that it is of sufficient intensity to induce a change in the chicks' affective state which persists in the subsequent test session, but is not intense enough to distract the chicks' attention away from foraging, and towards the buddy chicks behind the wire mesh. Indeed, as our earlier discussion suggests, a one minute period of social isolation might be sufficient to induce a change in the chicks' affective state without distracting the their attention away from foraging in the test session, in possible contrast to a longer period of social isolation, of three minutes, for example.

The experimental designs we have just discussed have involved manipulating the external environment of the chick. Alternatively, one could use a treatment which directly targets the physiological machinery of affect, for example by using psychopharmaceuticals. Such treatments are often well-validated (e.g. Feltenstein et al., 2004; Sufka et al., 2006; Warnick et al., 2006), and have advantages over environmental manipulations, in which any resulting change in an animal's behaviour can *arguably* be attributed to what the animal has learnt from those manipulations, without implying any role for current affect[174]. However, they may not always target affective state 'cleanly': e.g. they may also change activity levels, drowsiness, and so on (e.g. Feltenstein et al., 2004; although, of course, some of these behaviours might normally be expected to change as affective state changes), which could lead to considerable difficulty interpreting the results from a foraging-based task requiring locomotion around a test arena.

[174] For instance, if an animal is subject to environmental manipulations designed to induce an anxiety-like state, any resulting change in behaviour suggesting, for example, a greater tendency to attend to possible sources of threat, or to categorise ambiguous stimuli negatively, could arguably be due to a cognitively-based process of past experience shaping memory, schemas, attentional bias, risk sensitivity, probability assessments, and so on: I.e. without implicating current affect as a factor (Christine Nicol, personal communication).

To summarise, then, we did not find evidence to support our hypothesis that a treatment designed to induce a negative change in chicks' affective state would reduce their tendency to attack red food, a colour commonly found in aposematic animals, via a concurrent biasing of the cognitive processes involved in foraging decisions. We did, however, find evidence that the treatment changed aspects of the chicks' cognition, simply not in a manner we had predicted. The results suggested that following a three minute period of social isolation, a greater portion of chicks' attentional resources were occupied with re-establishing close social contact with conspecifics, compared to subjects who had received no such social isolation. This behaviour may be interpreted with reference to any change in affective state that occurred as a result of the treatment: i.e. it is addressing the source of that discomfort, and thus seeking to reduce it; but other interpretations which do not implicate affect are equally valid. Finally, the results provide an insight into how the cognitive mechanisms involved in foraging behaviour may be functionally structured, especially with regard to attention; this is interesting on its own terms, but also has implications for the design of foraging choice tests.

CHAPTER 6

GENERAL DISCUSSION

THESIS OVERVIEW

The opening chapter of this thesis discussed some of the problems associated with addressing consciousness within a scientific framework. These problems emerge in sharp relief in animal welfare science, since, for many people, concerns about animal welfare are contingent on the possibility that animals have subjective experiences. Through careful comparative work, a dossier of circumstantial evidence can be built in support (or otherwise) of that possibility. In parallel to such endeavours, however, there is a pressing need for animal welfare scientists to develop objective measures which are as faithful a correlate as possible to any subjective emotional experience an animal might be having. Many such proxy measures have been developed, and these have made a very valuable contribution to our understanding of the animals under our care. However, these indicators have some important limitations, which a recent approach, stressing affect-related changes in information-processing, aims to address (Paul et al., 2005).

The vast majority of the scientific literature pertaining to the relationship between emotion and information-processing has studied humans. Reviewing that research yields a number of hypotheses which can then, in principle, be tested in non-humans; many of these concern the characteristic biasing of certain cognitive processes – such as judgements, interpretations, memory and attention – across affect. Since the experimental work from which those hypotheses are derived relies heavily on verbal paradigms, there is therefore a need to develop novel, non-verbal tasks for use with non-human animals.

In the first experimental chapter, we reviewed one such study (Harding et al., 2004), which, employing an innovative paradigm, tested one such hypothesis, and yielded some encouraging results. We then adapted Harding et al's (2004) methodology in an alternative design intended to address some interpretational

issues originating from the unbalanced contingencies they employed. We applied this design in three experiments, described over the next three chapters (2-4), employing rats as experimental subjects. In the first two experiments, we attempted to manipulate the rats' affective state using unpredictable husbandry events, and changes in homecage 'enrichments', respectively. In the last of these (Chapter 4), we employed a treatment designed to vary subjects' level of food motivation, via pre-feeding, prior to their probe-tests. In our final experimental chapter (5) we explored a novel paradigm with domestic chicks, adapting an experimental design previously used to investigate the elicitation of colour-related foraging biases.

Our discussions, in each of these experimental chapters, have been reasonably extensive, and we shan't revisit these again in detail. However, we shall present some more general conclusions, and to allow these to be better informed, we will briefly review some other studies, recently published during the preparation of this thesis, which have been informed by a 'cognitive' approach to the understanding of non-human emotion.

OTHER RECENT STUDIES INVESTIGATING AFFECT-RELATED 'COGNITIVE BIAS' IN NON-HUMAN ANIMALS

Recently, Matheson et al (2008) also modified Harding et al's (2004) 'go/no-go' paradigm by employing a two-choice procedure, this time studying starlings. In their design, the contingencies were as follows: the *conditional stimulus (CS)* was the illumination of a light for one of two durations (2 seconds, or 10 seconds); the *conditional response (CR)* was a peck on either a flashing green lever, or a flashing red lever; and the *unconditional stimulus (US)* was the receipt of a food reward consistent in quantity, but differing in the latency with which it was dispatched (either 1 second, or 15 seconds).

Since starlings prefer shorter, over longer, delays to food (e.g. Bateson & Kacelnik, 1996, 1997), the authors predicted that subjects in a putative negative affective state would be more likely to respond in a 'pessimistic' manner (i.e. more likely to peck the key associated with the longer delay to food) when presented with probe

stimuli (in which the light was illuminated for 2, 3, 4, 5, 6, 7, 8, 9, or 10 seconds) compared to birds in a more positive affective state.

They attempted to manipulate affect by adjusting aspects of the birds' husbandry regime: an 'enriched' treatment consisted of a larger homecage, more 'environmental enrichments', less disturbance during cage-cleaning, and more predictable access to a water bath, compared to the alternative, 'standard', treatment they employed; the former was designed to induce a positive change in affect, the latter a negative change. Six starlings were trained to criterion and completed both subsequent treatments in a counterbalanced order (i.e. in a repeated-measures design). Each treatment lasted fourteen days, with probe-test sessions conducted once daily from days 5 - 14.

They conducted probit analyses, across probe value, on the birds' response choice, and then compared aspects of this fitted function in repeated-measures models. Overall, they did not find a significant difference between the two treatments in the probe value at which the probability of making either response was 0.5 (i.e. the 'bias'), nor in responding to the stimulus associated with the 'quick' (short-delay) reward. However, they *did* find a significant difference, in the hypothesised direction, in response choice when presented with the stimulus associated with the 'slow' (long-delay) reward ($p=0.040$). Otherwise, they found a tendency ($p=0.053$) for the middle section of the fitted function to be steeper for the 'standard' treatment group (i.e. for them to be more 'sensitive' to probe value).

Since the steepness of the fitted curve differed depending on which contingency group the starlings were in, indicating that the birds were less sensitive to probe value when in the 2-sec *(CS)* / 1-sec *(US)* (i.e. 'short=instant') group, the authors re-analysed the data from the more 'sensitive', 'long=instant' group only. This analysis found a significant difference in 'bias', in the hypothesised direction across treatment, and additionally found that 'sensitivity' was significantly higher for the 'standard' treatment group (i.e. that group were more discriminating across probe value). Their results, then, provided some very encouraging support for their hypotheses, and furthermore, by studying a different species, and employing a

different experimental design, they were able to investigate to what extent Harding et al's (2004) findings generalised.

Interestingly, as in a number of the experiments conducted in this thesis, Matheson et al (2008) found that contingency (namely whether the 2 second, or 10 second, light was associated with the short, or long, delay to food) interacted with the *US* to produce differing psychometric functions. As the authors note, the extent of that interaction may have been accentuated, or determined, by the selection of conditional and unconditional stimuli which varied the same parameter (time) to characterise their differences (i.e. birds 'waited' variously for (2 or 10 seconds) + (1 or 15 seconds) for food reward: i.e. for 3, 11, 17, or 25 seconds). As the authors noted, since delay to food was the *a priori* determinant of *US* value, it may have been more advisable to employ a different *CS*, such as different tonal frequencies.

In addition, Matheson et al (2008) further cite evidence suggesting that if there is a longer delay before a reward is obtained, then that reward may be valued more highly than a reward received at lower cost (e.g. Kacelnik & Marsh, 2002; Klein et al., 2005), an observation which has some interesting implications for the direction of hypotheses based on the *a priori* identification of preferred outcomes. To complicate matters a little further, perception of the passage of time itself may be biased across affect, at least in humans; for example, people with depression (e.g. Blewett, 1992; Bschor et al., 2004; Mundt et al., 1998) or in anxious or fearful states (e.g. Langer et al., 1961; Watts & Sharrock, 1984) may be more likely to overestimate the duration of temporal intervals, perhaps reflecting differences in the deployment of attentional resources, and/or levels of arousal (e.g. Droit-Volet & Meck, 2007; Wittmann & Paulus, 2008).

In another recent study with starlings, Bateson & Matheson (2007) trained 6 birds to flip a lid from the top of a petri dish to access mealworm prey beneath. Palatable mealworms were covered with a white lid, whilst unpalatable mealworms (injected with quinine) were covered with a darker-coloured lid (printed in 80% greyscale). Once the birds were flipping reliably more white lids than darker lids, they entered two sequential treatments. The starlings' homecages in the 'standard' treatment remained as they had been during training, whereas their homecages in

the 'enriched' treatment received additional 'enrichments'; two birds received the former treatment first, whilst four received the latter treatment first. Once the birds had been in their respective treatments for two days, they received a number of probe trials, over the next five days, in which they were presented with 'probe' lids of intermediate shades (20%, 40% & 60% greyscale), none of which covered mealworms of any sort.

Their results indicated that the 'enriched' group flipped more of the 20% lids than the 'standard' group (i.e. were more 'optimistic' with regard to the 'probe' lid closest in tone to the palatable mealworm reference (white) lid), but only for the counterbalanced group receiving the 'enriched' treatment first. If we assume that the difference in the tendency to flip this lid, across treatment, reflected a change in affect as hypothesised, it's possible that moving from an 'enriched' to a 'standard' cage produced a comparatively more negative affective state in the latter ('standard') treatment, compared to birds who had known, in the recent past, no better (i.e. who received the 'standard', or control, treatment first). So, whilst as a go/no-go design, the task they employed bears some similarity to the design adopted by Harding et al (2004), and thus may be relatively more vulnerable than some alternative designs to certain confounds (as discussed in Chapter 2), their results were again encouraging, and it was another novel addition to the increasing canon of tasks informed by a 'cognitive-bias' approach.

Returning to rodent-based studies, Burman et al (2008b) manipulated 'environmental enrichment' in an experimental design which employed different spatial locations as probes. They provided twelve rats with extra homecage 'enrichments', above and beyond their pre-existing standard fare, and removed some of the pre-existing 'enrichments' from the homecages of a further twelve rats. They then trained all 24 rats, in a test arena, to run from a starting-box to a pot. When the pot was in one location (e.g. far to the right), it contained an accessible food reward, but when it was in an alternative location (e.g. far to the left, counterbalanced across subjects), the rats found no accessible reward inside it. Once the rats were reliably running more quickly to the baited pot than they were to the un-baited pot, they received a number of probe trials, in which the pot appeared in one of three locations intermediate to the 'reference locations' with

which they had previously been trained. The experimenters found that whilst running latencies did not differ between the treatment groups at each of the 'reference' locations, it *did* differ, in the hypothesised direction, for the probe location closest to that occupied by the unrewarded 'reference' pot (latencies did not differ, across treatment, with respect to any of the other probe locations).

Furthermore, employing a somewhat similar treatment, but a different paradigm, in another recent experiment Burman et al (2008a) removed a number of pre-existing 'enrichments' from the homecages of twelve rats (the 'unenriched' group), leaving the homecage environment of a further twelve rats unchanged ('enriched' group; i.e. with pre-existing 'enrichments' intact). They then trained the rats to run down a runway to receive either 1 or 12 pellets of food (with each rat always receiving the same amount of food, and with this quantity counterbalanced across treatment). Once the rats receiving 12 pellets were running reliably faster than the rats receiving 1 pellet, *all* the rats thereafter received just 1 pellet at the end of the runway. The group 'down-sized' from 12 to 1 pellets exhibited the characteristic behavioural change associated with such 'successive negative contrast' (e.g. Flaherty, 1996): namely, they ran, on average, more *slowly* than the rats who had *always* received 1 pellet of food. However, the slower latency of the 'unenriched' rats persisted significantly longer than the slower latency of the 'enriched' rats, before finally recovering to match that of the other experimental groups. Here, the mechanisms underlying the observed treatment-related difference may differ from those implicated in a number of the other 'cognitive bias' tasks reviewed above (and also those conducted in this thesis), suggesting an affect-related bias in the perceived impact, or significance, of reward withdrawal, and/or of mnemonic processes in which negative memories, of downsizing, remain more salient, for a longer period of time.

METHODOGICAL AND THEORETICAL ISSUES IN THE STUDY OF AFFECT-RELATED 'COGNITIVE BIAS' IN NON-HUMAN ANIMALS

So, as this brief review illustrates, an 'information-processing' (or 'cognitive') approach to understanding affective states in non-human animals is beginning to show considerable promise: whilst underlying mechanisms remain to be elicited, a

variety of paradigms, studying a variety of species, have yielded encouraging results. To ensure good progress continues, though, what conclusions can we draw from the present state of such research (including the experiments conducted in this thesis), and what cautionary correctives might we now need to consider?

One proposed benefit of a 'cognitive' approach is the comparative clarity of *a priori* predictions (e.g. Burman et al., 2008b). This contrast is made in comparison to other, more 'traditional' approaches to the behavioural measurement of affect, including the elevated plus maze and open field, in which hypothesised changes in behaviour are not always clear (as we discussed, for example, in Chapter 2).

However, a number of the studies, and discussions, conducted in this thesis suggest such clarity cannot be taken for granted, encountering, as we have, counter-intuitive results, and a multiplicity of predictions when re-visiting the human-based literature. It's important to note that this isn't a negative development: on the contrary, by acknowledging such complexity we have gained a more rounded view of our results, and have gone at least part way to accounting for some curious findings, which might otherwise be discounted as anomalous noise. We will now consider the factors which might contribute to the complexity of *a priori* predictions in tests of affect-related non-human cognitive bias.

Clearly, the treatments themselves may be an important contributor to the nebulous nature of any predictions, and we discussed the pros and cons of a variety of methods of manipulating affect in Chapter 5. Perhaps the clearest violation of our *a priori* assumptions occurred in our opening experiment (Chapter 2), in which the unpredictable housing treatment appeared to induce a counter-intuitive change in putative affect, together with an 'optimisation' of cognitive performance. Interestingly, the unpredictable husbandry regime employed by Matheson et al (2008) was similarly associated with an improvement in performance: in terms of maximising food return through improved accuracy in a relatively demanding cognitive-behavioural task. Here, starlings in the 'standard' treatment (*cf* 'enriched') were more sensitive (i.e. more discriminating) across *probe value,* with a (near-significant) tendency to be more accurate with regard to the 'long delay reward', without any detriment to their accuracy with regard to the

opposing reference stimulus. The authors interpreted this by reference to 'depressive realism': the notion that the expectations, and judgements, of depressed humans are closer to the objective truth than that of non-depressed people (e.g. Taylor & Brown, 1988). However, as discussed in our opening chapter, some have argued that 'depressive realism' is empirically-equivocal and situation-specific (e.g. Power, 1999); in certain scenarios, depressed people, for example, appear *less* realistic than non-depressed (e.g. Moore & Fresco, 2007). Alternatively, it's possible that the difference, in accuracy and discrimination, Matheson et al (2008) uncovered between their treatment groups may have reflected a stress/arousal-facilitated increase in cognitive performance, as perhaps we encountered in Chapter 2. These observations suggest that for treatments employing unpredictable interventions, the scope for predictions to be mispecified is relatively large; this may, in part, reflect the diversity of methods adopted when attempting to manipulate affect in this manner: e.g. compare the comparatively mild interventions employed in this thesis, to those adopted by Willner et al (1987).

Changes in 'enrichment', on the other hand, and in particular the removal of pre-existing homecage substrates, *appear* to be more reliably associated with changes in behaviour in keeping with *a priori* predictions relating to affect-related biases in information-processing (e.g. Bateson & Matheson, 2007; Burman et al., 2008a; Burman et al., 2008b). Indeed, there was some suggestion, in Chapter 3, that the same might be true in this thesis, and we noted, in the corresponding discussion, that such evidence (of an 'optimistic bias') is contingent on being able to satisfactorily model both *accuracy* (in terms of responding to the reference tones), and *responses to ambiguity*. However, attempting to distil out the latter may not be a simple matter. In discrimination tasks, for example, animals rarely perform 'perfectly' (i.e. the criterion at which an animal is considered 'trained' is generally set some way below 100%). More generally, the (highly successful; e.g. Wickens, 2002) application of signal detection theory in psychophysics is premised on the assumption that environmental, or cognitive 'noise' routinely obscures signal clarity (e.g. Atkinson et al., 1996). In the case of the operant paradigm adopted in this thesis, for example, it may not be appropriate to conceptualise the reference trials as entirely 'unambiguous'.

What about the probe stimuli? *A priori,* their level of ambiguity may depend, in part, on their proximity to the reference stimuli (and *vice versa:* i.e. the 'ambiguity' of the latter may depend, in part, on their proximity to the former), but how? Are those probe values furthest from the reference stimuli the most, or least, ambiguous? For example, are they perceived as being *clearly neither* of the reference tones, or are they on the cusp of being categorised as *either*? Might some probe values appear even less 'ambiguous' than the reference stimuli themselves? The phenomenon of peak shift (i.e. peak responding, in discrimination tasks, to probes rather than reference stimuli; e.g. Ghirlanda & Enquist, 2003) suggests that, in some circumstances, this might be so.

These are important questions, since constructs such as *effectiveness, accuracy* and *efficiency* need to be distinguished from *'optimism'* and *'pessimism'*. Otherwise we risk confounding affect-related changes in cognitive *capacity,* with affect-related changes in cognitive *selectivity* (e.g. Dalgleish, 2003). To this end, the designs employed by Burman et al (2008a; 2008b), in which the training, as well as testing, were conducted whilst the subjects were in their respective treatment groups, controlled, at least in part, for confounds relating to cognitive *capacity*. In each instance, such confounds (e.g. in learning ability) were not found, but if they had been, it might have been possible to quantify them (e.g. as 'number of trials until criterion reached'), and then model them as a covariate. Such a design may not always be desirable, however: animals in a putatively more negative affective state may learn equally 'well' (e.g. quickly), yet the 'contents' of that learnt representation might differ from that of subjects in more positive states; such conceptual differences *could* have important consequences for the interpretation of behavioural responding in any subsequent tests. On a more ethical note, if the training phase is long (e.g. if the subjects take some time to learn what might be a difficult task), the animals' welfare might be too greatly compromised if they are kept in particular treatment groups throughout that process.

Otherwise, the clarity, or otherwise, of *a priori* predictions in tests of 'cognitive bias', may, in part, be determined by the choice of unconditional stimuli. In Chapter 2, for example, our findings didn't support our hypothesis that an unpredictable housing

treatment would induce a pessimistic style of responding, whereas, in Harding et al's (2004) study, it did. In the case of the former, the unconditional stimuli were either one, or two pellets of food, whereas in the case of the latter, they were either one pellet of food, or no food, the delivery of white noise, and a longer delay until the next trial. Assuming, for the sake of argument, that such unpredictable husbandry treatments induce a greater tendency to interpret ambiguous probe stimuli as signifying a negative outcome, which of these various outcomes is the most negative? If the receipt of food *per se,* is better than nothing, yet nothing is better than the delivery of white noise, then that might account for the differences in results. On the other hand, might the negative consequences of *inaccurately* judging a probe stimulus as heralding the delivery of two pellets of food be deceptively great, because the quantitative difference here, between two, and zero, pellets of food, is particularly large? Since, as we discussed in Chapter 1, there is some evidence that subjective probability assessments, with respect to positive and negative events, differ across types of negative affect (e.g. depression, and anxiety), such issues may be an important determinant of success when searching for affect-related biases. But they are conceptually-challenging: e.g. if the delivery of two pellets of food is better than the delivery of one pellet of food, yet the delivery of both is better than nothing, is the delivery of one pellet actually *negative,* or just *less positive*?

On a different tack, it's possible, of course, to dispense with conditional stimuli, and study animals' responses to probes which have an unlearnt affective significance (e.g. predator-like stimuli, such as models (e.g. Jones et al., 2007), or something more elemental, such as fast swooping movements, and so on), varying the resemblance of the stimuli to the predator (for instance) it's intended to mimic. Such designs could potentially dispense with a considerable amount of training, but of course are prone to habituation, at least if repeated over a number of trials (i.e. there may be only a small window of opportunity to observe responses, and perhaps such a paradigm may only be useful if the effect size was very large). Otherwise, there may be affect-related biases in non-human animals which are, *a*

priori, less obvious candidates for enquiry,[175] but may suggest themselves to us through careful consideration of the adaptive significance of various affective states.

None of the studies reviewed so far have employed strongly aversive outcomes as unconditional stimuli. This is despite the fact that the evidence for affect-related biases in the processing of threat is very strong in humans, as we discussed in Chapter 1. It's also despite the fact that anxiety, strongly implicated in such biases, is perhaps, *a priori,* more likely than depression to be found in a range of species, because its functional significance, at least when expressed non-clinically, is more obvious. Since the experiments we've been discussing have approached the issue of affect-related cognitive bias from an animal welfare perspective, it's not surprising that strongly aversive unconditional stimuli have not been employed. However, approaching the issue from an alternative perspective, namely that of affective neuroscience, a recent paper *has* investigated interpretive biases relating to ambiguous threat.

Tsetsenis et al (2007) studied knockout mice lacking an active serotonin receptor gene, the dysfunction of which has been implicated in anxiety and depression. On the first day of fear conditioning, they presented both the knockout, and wild-type mice, with five trials; on the first, third and last of these trials, a 3kHz tone was presented for 20 seconds, terminating in the presentation of a light for 20 seconds, which itself terminated in a foot shock. In the second and penultimate trials, the mice were presented with the 20-second 3kHz tone alone; hence, the light always heralded an aversive shock, whereas the 3kHz tone heralded a light, then a shock, on 60% of occasions, and neither of these things on 40% of occasions.

The next day, the mice were presented with the 3kHz tone alone for 6 minutes, and the light alone for 6 minutes. All mice exhibited high levels of freezing behaviour to the light, and low levels of freezing behaviour at 'baseline' (i.e. when neither the light nor the tone were presented); however, the knockout mice froze significantly

[175] *cf* attentional biases for upper or lower regions of physical space (e.g. Meier & Robinson, 2006), and judgemental / perceptual biases with regard to darkness and light (e.g. Meier et al., 2007), across affect in humans.

more in response to the tone than the wild-type mice. This was interpreted as analogous to the readiness with which some anxious people respond to 'ambiguous' cues (akin to the tone, in this instance) with an inappropriately-large fear response, often to the detriment of their quality of life and general functioning.

Clearly, this was an interesting study, and is perhaps itself analogous to experimental paradigms of fear extinction and recall, in which a neutral stimulus (e.g. a light) is paired with an aversive one (e.g. a foot shock) a number of times, such that the light itself (i.e. without presentation of the shock) induces fear-related behaviour (e.g. freezing). This behaviour abates as the light is then repeatedly presented by itself (i.e. without any associated shock, inducing 'extinction'). When *later* presented with the light (alone), however, some animals, under some conditions, show impaired extinction recall: e.g. they freeze (as reviewed in Milad et al., 2007, for example). As Bishop (2007) notes, the light has become, "in effect, an emotionally ambiguous stimulus linked to representations of both threat (from acquisition) and safety (from extinction)". Given the substantial insights which can, potentially, be gained from such an experiment, should animal welfare scientists, developing proxy indicators of affect, also consider employing such aversive stimuli? We shall consider this issue later, but first we'll further widen our discussion of the factors which may contribute to the equivalence of *a priori* predictions in 'cognitive bias' tasks.

Matheson et al (2008) noted that similarities between their findings, and that of Harding et al (2004), suggested that animals in a poorer environment may be more averse to relatively bad outcomes, a conclusion which has some adaptive appeal. However, as we briefly mentioned in our opening chapter, a number of studies have found that humans are more sensitive to losses (and more risk-averse in their decisions) when in a more *positive* affective state (e.g. Kliger & Levy, 2003; Nygren et al., 1996). Here, then, we have a potential problem for many of the studies reviewed above, and, of course, for many of the studies presented in this thesis. If, for example, an animal in a more negative affective state is more risk-prone than an animal in a more positive affective state, then even if the former subject estimates the likelihood of a positive outcome as lower, or the likelihood of a negative outcome as higher, than the latter subject (as we might predict from the

human literature), an animal in a more negative affective state may nevertheless be *more likely* to respond as if expecting the better outcome. Operationally, therefore, such an animal may be more optimistic, yet in their cognitive judgements of the likelihood of various occurrences, they may, in fact, be more pessimistic.

In Chapter 1, we encountered a possible way to address such a problematic issue. If affect-related risk-proneness/aversion only manifests when the multitude of potential outcomes is explicit: e.g. in human terms, when one is weighing up the likelihood of various outcomes, and selecting among several possible choices when making one's decision, then it's perhaps best conceived as a *response bias* (as we defined it in our opening chapter). If, however, a (human, or non-human) animal's response to ambiguity instead reflects what we termed, in that chapter, an *interpretive bias* – i.e. only one semantic representation is processed, and acted upon – then there is, by definition, nothing to risk. The resulting behaviour of such animals may therefore be a more faithful correlate of a genuinely-'optimistic' bias in information-processing.

So, how might we put such an observation into practice? As discussed in the opening chapter, such issues have been addressed, in human studies, through the employment of priming techniques (e.g. Calvo et al., 1994; Macleod & Cohen, 1993; Richards & French, 1992): i.e. facilitated (or otherwise) responding to stimuli following the presentation of an ambiguous prime, one interpretation of which is congruent with the correct identification of the target stimulus (e.g. correctly identifying it as a 'real' word, rather than a non-word). In these experiments, facilitation is operationalised by response latency across response choice. The operant discrimination tasks conducted in this thesis have been, to some extent, analogous to such a procedure, with the ambiguous tone 'priming' the subject to respond one way or the other; in such designs, response latency *may*, therefore, be a reasonable proxy of such facilitation.

A NOTE ON ETHICS

There's been relatively little discussion, so far in this thesis, of the ethical implications of the experiments we have conducted, nor of the hundreds of

comparable studies elsewhere. In keeping with the relatively dispassionate protocol of scientific prose, we've talked about the treatments we've employed in terms of their scientific merit, rather than engaging in a more expansive debate of the costs, in potential suffering, to the animals concerned; nor indeed have we given a more personal account of any discomfort experienced by those involved in such experimentation. However, we'd like to make a few points on such matters now.

Firstly, by the very nature of the objectives they're trying to achieve, many treatments designed to induce a change in an animal's affective state clearly *do* have potential costs (of suffering) to the animals concerned. More generally, keeping animals in the confined, perhaps unwittingly ill-prescribed, conditions of a laboratory has potential welfare costs *per se*. The onus is (of course) on those of us engaged in such research to constantly seek to refine the procedures, and general conditions, employed. Some of these refinements (or replacements) might be counter-intuitive: for example, an (invasive) psychopharmaceutical treatment (e.g. via injections), as discussed in Chapter 5, might compromise welfare *comparatively* less than alternative options if its effects are very transient, or, most obviously, if its intended effects are to induce a positive change in affect. Indeed, there may well be merit in adopting treatments designed to induce a positive change in affect *per se*, but *even those* have implicit costs: contrasted, as they are, against a control group we *know* 'could have it better'.

More generally, there is an understandable, and difficult, tension between employing treatments we think are likely to be 'mild', or even positive, in their effects, and their ability, contrasted against alternative procedures, to definitively answer the scientific questions we are posing. For example, the (scientifically) impressive experiments conducted by Tsetsenis et al (2007), reviewed above, employed a mouse strain genetically-engineered to be anxious and depressed, or at least engineered to express some of the neurological characteristics associated with such affective states (Nader & Balleine, 2007). The ethical implications of employing a strain which is, arguably, *designed to suffer,* are profound, yet the scientific progress they made (using aversive foot-shock stimuli) is, potentially at least, substantial (e.g. Nader & Balleine, 2007).

In contrast, in this thesis we have employed treatments which are designed to be milder than the majority employed elsewhere, perhaps simulating aspects of a (very negligent) husbandry regime; even so, society's notion of what is acceptable is constantly, and for the most part progressively, refined, and it may be that in years to come even such 'mild' manipulations will no longer be permissible. More generally, such an observation alerts us (more specifically: me) to the importance of constantly questioning the ethical implications and scientific merits of our work, and of guarding against finding comfort in the fact that we presently adhere to societal norms.

As this discussion suggests, these issues are, of course, very challenging ones for the scientist to address, and charting a course through them is not an easy matter. As a general point, though, by explicitly discussing them, we can at least contribute in a small way to an ethical debate which aspires to be honest and intelligent; a necessary corrective to a scientific culture increasingly driven by coarse indices of achievement (e.g. Lawrence, 2007).

CONCLUDING REMARKS

What's different about 'cognitive bias tests'? A number of the papers we've briefly reviewed in this chapter have used that label to describe the test they've employed (e.g. Burman et al., 2008b; Matheson et al., 2008), and whilst that might be a very reasonable description, it somewhat suggests they are sensitive to processes fundamentally different from a number of other paradigms. It's an obvious point, but perhaps worth making, that this is unlikely to be the case, at least not in all instances. For example, one of the key measures in an elevated plus maze (EPM) is the proportion of time spent in the closed arms; this is generally hypothesised to be greater when anxious. It's conceivable, in such circumstances, that an 'anxious' animal interprets the emotional ambiguity of the open, exposed arms as more likely to be threat-related: e.g. perhaps more likely to be associated with attack. Hence rather than risk an outcome which may be subjectively more probable (than if the animal were less anxious), the subject stays in the least-exposed areas; similar inferences could be made with regard to other behavioural measures, such as the open field, and novel object test.

The mechanisms implicated in the 'cognitive bias' studies we've discussed may differ in many ways, both from each other, and from 'traditional' behavioural tests of animal affect. It's not *necessarily* the case that a 'cognitive bias' test is investigating mechanisms other than those examined in the EPM, etc., but for such a research programme to progress, various functional and mechanistic hypotheses need to be distilled, tested and compared (cf risk-sensitive foraging in non-human animals; e.g. Kacelnik & Bateson, 1996; Kacelnik & Bateson, 1997). As a *general* observation, though, in adopting a 'cognitive approach' to emotion, a diverse range of useful novel methods have been developed, in a short timescale. What makes this approach so fertile? Through an explicit review of the large human-based literature concerning information-processing and affect, a range of hypotheses have been formulated for investigation in non-human animals. Since these centre on (superficially) substrate-neutral observations concerning causality and function, this allows flexibility, and transferability, of experimental design; this is the likely key which will determine the future utility of such an approach.

There are many examples of such flexibility. In a number of the paradigms reviewed in this thesis (e.g. Bateson & Matheson, 2007; Burman et al., 2008b; Harding et al., 2004; Matheson et al., 2008; Tsetsenis et al., 2007), the *valence* of the potential outcomes associated with an 'ambiguous' stimulus, for instance, can be manipulated between those which are aversive, such as potential attack, pain, or unpalatable food, and those which are positively-valenced, such as palatable food. In addition, as we discussed in Chapter 1, by varying the relevance of the unconditional stimuli to the experimental subjects, we may be able to distil specific sources of fear (e.g. Teachman et al., 2008), for example, or relate them to more motivational constructs, such as hunger, or thirst (e.g. Changizi & Hall, 2001). Furthermore, 'ambiguity' can be quantified, at least physically (e.g. in terms of kHz value, stimulus duration, etc.), between known outcomes, i.e. through reinforcement schedules, as manipulated by the experimenter. Furthermore, the experimental conditions can be manipulated in a manner which renders them less aversive to experimental subjects. Rather than a single enforced confrontation with a novel, perhaps unpromising situation, for example, there is the opportunity for more extensive data to be built up over a longer period of time, in a situation which need not be aversive at all to the animal (in fact, may even be 'enriching').

Our discussion, above, suggests some correctives to current terminology may be needed so that 'cognitive bias' tasks are placed in a framework which includes other, more 'traditional' behavioural measures of non-human affect. However, the flexibility an 'information-processing', or 'cognitive', approach allows, can have real scientific, and ethical, benefits.

More generally, given the richness of the literature pertaining to the relationship between information-processing and affect, the scope for developing further novel tests of non-human emotion, and to modify existing tests, is great. Our understanding of animal cognition is rapidly maturing, and by cross-referencing the predominantly human-based literature on information-processing and affect, with the animal-based literature concerning cognition, there is substantial scope for further progress. The simple modification of an existing foraging task in chicks, outlined in Chapter 5, illustrates this. That study also suggested that such research, and the results it yields, may prove more complex than initially anticipated, yet that, in turn, can be turned into fertile ground for future scientific investigation.

A 'cognitive' approach to the understanding of non-human affect will also benefit from further integration with fields outside of animal welfare science: for example, behavioural ecology, psychopharmacology, affective neuroscience, and so on. In particular, there perhaps needs to be more explicit discussion of the functional significance of various affective states, and their proposed correlational relationship with cognitive bias. After all, historical attempts to study, and elicit, human abilities and behaviours in non-humans, perhaps most famously aspects of human language, have suffered from an anthropomorphic bias which has ignored evolutionary, functional, and anatomical considerations (e.g. Pinker, 1994).

Nevertheless, the mechanisms we have studied in this thesis, are, *a priori,* perhaps likely to be found in a large number of different species, since they involve state-dependent changes in the processing of fitness-relevant stimuli. That simple observation, in turn, opens up the possibility that affect-related mechanisms may be implicated in a great range of phenomena documented in non-human animals, such as changes in vigilance across predator type and density, state-dependent

risk-sensitivity, response to aposematic warning colouration, and so on. Through such cross-disciplinary endeavour, the scene is set, then, for the science of non-human emotion to grow in exciting, and unanticipated, directions.

APPENDICES

APPENDIX A: PILOT STUDY - RATS' PREFERENCE BETWEEN TWO RESPONSE OPTIONS REINFORCED WITH TWO DIFFERENT QUANTITIES OF FOOD

Introduction

The present study was designed to test whether reinforcement schedules of 1, and 2, pellets of food sufficiently differ from each other, in terms of putative hedonistic value, to satisfy a basic condition of their use in a subsequent 2-choice experiment with rats.

Method

Subjects and housing

The study subjects were 12 (6 male; 6 female) Lister-hooded rats (*Rattus norvegicus*; Harlan UK Ltd., Bicester, UK), purchased at 3 months of age. They had previously been subjects in a choice-chamber experiment investigating preference for variable visual environments, and were 6 months old at the start of the present study.

The rats were housed in stable same-sex groups of three, in cages measuring 56cm (L) x 34cm (W) x 19cm (H), with a 12:12 hour lights on/off regime (lights off at 11am). The cages contained sawdust bedding (Lignocel), shredded paper for nesting, and an enrichment toy. They had ad libitum access to food (Eurodent Diet 22%; LabDiet, Richmond, IN, USA) and water.

Apparatus

Two operant chambers, of identical design, were used, in two different experimental rooms. The operant chamber measured 52cm (L) x 30cm (W) x 35cm (H). Three of the walls and the floor were metal (the latter was covered with 1 litre of sawdust bedding (Lignocel)), the long rear wall was Perspex, and wire mesh covered the ceiling. A food hopper was located centrally on the long metal wall (3.5cm above the floor), with a retractable lever on either side (4cm away from the side of the trough, 8cm above the floor). The chamber was illuminated with a 1.12W white light bulb. A water bottle hung at the rear of the chamber. The house light, levers and pellet dispenser were manufactured by Coulbourn Instruments (Allentown, PA, USA), and were operated using their Winlinc software. The operant chamber and food hopper were custom-made. The hopper delivered Bioserv (Frenchtown, NJ, USA) Dustless Precision Pellets (45mg).

Procedure

All training and testing was conducted in the rats' dark phase. Each rat received one session per day, 5 days per week (Mon-Fri), always in the same experimental room. The rats were trained and tested in the same order each day. Lever position (left or right) and quantity of reinforcement (one or two pellets of food) contingencies were counterbalanced as far as possible across sex of subject, experimental room assignment, and order of training and testing.

The chamber, including the hopper and levers, was sprayed with 70% ethanol solution and then wiped dry, and the floor was covered with fresh sawdust, prior to each rat's session. For a given rat, the two levers were physically swapped between sessions to control for any small, unintentional differences in design (e.g. lever length, force needed to depress the lever, etc.)

All rats received 5 sessions of magazine training. Each session started with the presentation of one lever (left or right position counterbalanced across CR-US contingencies, and alternated between sessions for a given rat) which, if pressed, resulted in the immediate delivery of one food pellet. If 10 lever presses were

made, the lever was retracted, and the other lever was presented and reinforced on the same schedule. This alternating pattern continued until the rat had made 60 lever presses, or 60 minutes had elapsed, whichever first. In addition, regardless of any lever-pressing, a parallel schedule was in operation in which the active lever retracted for one second prior to the automatic delivery of one food pellet. For the first three sessions, this autoshaping procedure occurred every minute, for the final two sessions it occurred every 3 minutes. This magazine training procedure was based on Mattel & Meck (1999).

In the final (5^{th}) session of magazine-training, all rats made 60 lever-presses bar one. This male rat received two additional days of magazine-training, the first of which included hand-shaping, but still failed to make 60 lever-presses in the second session, and was excluded from the remainder of the experiment.

Following magazine training, each rat underwent a series of quantity preference test sessions. Lever presses were now differentially reinforced: pressing one lever (e.g. the left) always delivered 2 pellets, pressing the other lever (e.g. the right) delivered 1 pellet (these positions were counterbalanced across rats, but the contingencies were the same across all sessions for each rat). The trials were arranged into 4 identical blocks in each session: 2 forced-choice trials followed by 10 free-choice trials.[176] In a given session, the order of forced-choice trials alternated, so that presentation of the left lever was followed by presentation of the right in the next forced-choice trial. The identity of the lever in the first forced-choice trial was counterbalanced across rats, and for a given rat this alternated between sessions. In free-choice trials, both levers were presented. Following a lever press, all presented levers were retracted, and the appropriate number of food pellets was dispensed. Each session terminated after 48 lever presses had been made, or 60 minutes had elapsed, whichever came first. The inter-trial interval (ITI) was 60 seconds.

---

[176] The forced-choice trials, in which only one lever was presented, ensured the subjects were regularly exposed to both contingencies, and were additionally designed to counter any side preferences which may have emerged regardless of the reinforcement regime (Abeyesinghe et al., 2005).

Each rat received 8 of these sessions, and then a further 8 sessions in which the number of presses required to receive reinforcement on the 2-pellet lever increased from session-to-session in the following increments: 2,3,4,5,6,7,8,13.

Results

Preference test: food quantity

All rats completed 48 trials in each session, except for one rat who completed 47 on the first day (48 thereafter). Due to equipment malfunction, data for only 10 rats were available on Day 3 (n=11 for all other days).

Figure 7.1 shows the mean percentage of free-choice trials in which a press on the 2-pellet lever was made in each session. For each day a one-sample t-test was conducted on these means. Table 7.1 shows that the percentage of presses on the 2-pellet lever was significantly greater than 50% chance from Day 4 onwards.

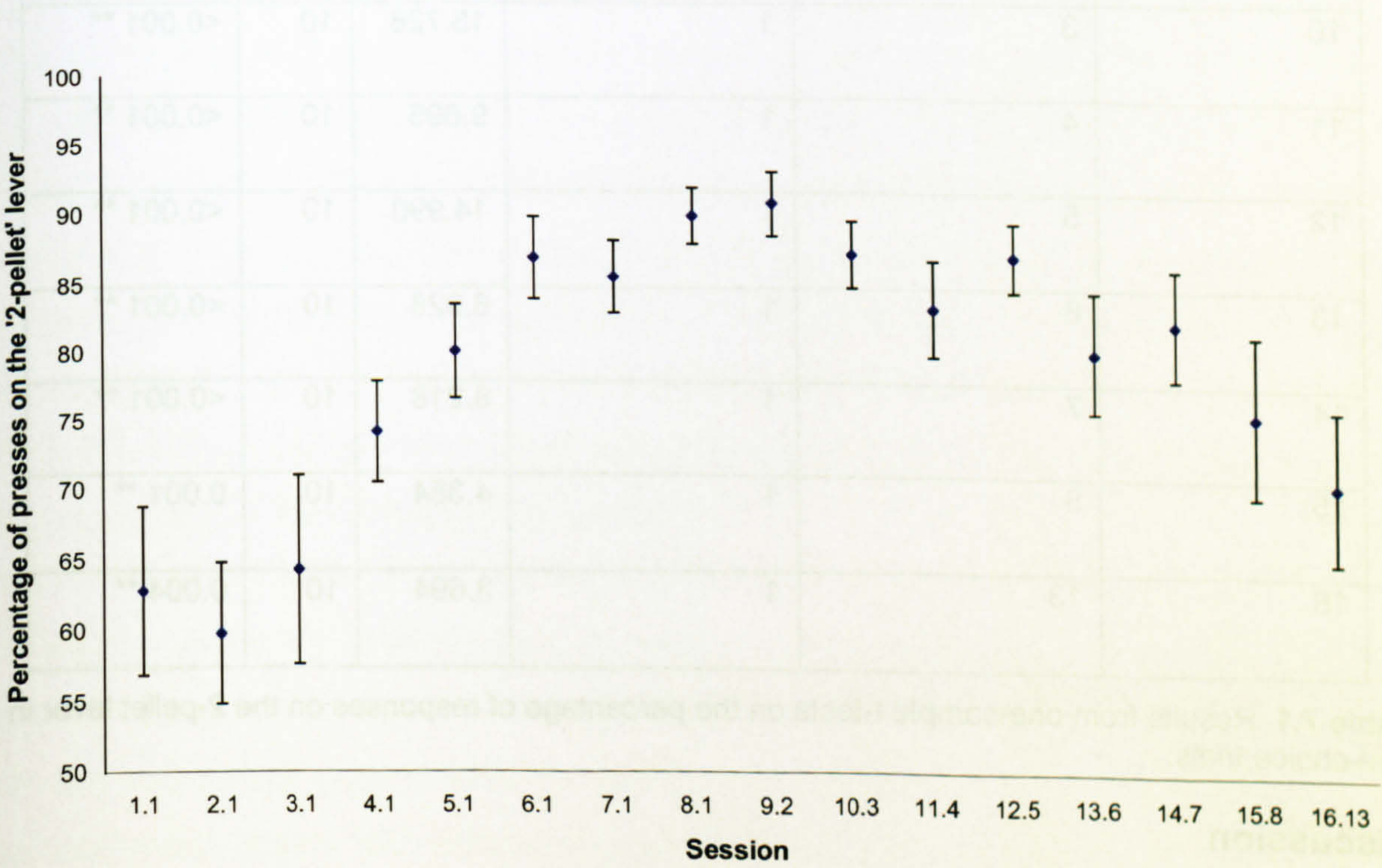

Figure 7.1 Mean percentage of responses on the 2-pellet lever in free-choice trials (+/- 1SEM). The Session names on the x-axis take the format of 'Session number: fixed ratio on 2-pellet lever.'

Day	Fixed ratio schedule		t	df	p-value (2-tailed)
	2-pellet lever	1-pellet lever			
1	1	1	2.146	10	0.057
2	1	1	1.961	10	0.078
3	1	1	2.137	9	0.061
4	1	1	6.571	10	<0.001 **
5	1	1	8.644	10	<0.001 **
6	1	1	12.604	10	<0.001 **
7	1	1	13.832	10	<0.001 **
8	1	1	19.815	10	<0.001 **
9	2	1	17.402	10	<0.001 **
10	3	1	15.728	10	<0.001 **
11	4	1	9.695	10	<0.001 **
12	5	1	14.990	10	<0.001 **
13	6	1	6.928	10	<0.001 **
14	7	1	8.218	10	<0.001 **
15	8	1	4.354	10	0.001 **
16	13	1	3.694	10	0.004 **

Table 7.1 Results from one-sample t-tests on the percentage of responses on the 2-pellet lever in free-choice trials.

Discussion

The results indicate that, when given a free choice, the rats chose a response which yielded two pellets of food significantly more often than a response (of equal cost) that yielded just one. This preference persisted after the cost (in terms of number of lever presses) of making the 2-pellet response increased 13-fold.

The present study was designed to test whether these unconditioned stimuli significantly differ from each other (in terms of putative hedonistic value) to satisfy a basic condition of their use in a subsequent 2-choice experiment, nominally measuring cognitive bias. This basic premise has therefore been satisfied.

APPENDIX B: DERIVING THE STANDARDISED SCALE OF SINGLE-FREQUENCY PROBE STIMULI

For the contingency group trained to associate responding to 4kHz with delivery of two pellets of food (*4kHz=2pell*), the auditory stimuli presented in the single-frequency probe test sessions were log-transformed to create a standardised scale using the following formula (Bill Browne, personal communication):

$$(\log(x) - \log(a) / \log(b) - \log(a)) + 1$$

where: *a* = 2000Hz; *b* = 4000Hz; *x* = the tonal frequency to be transformed.

By adding a constant of "1", the resulting scale was easier to interpret with regard to reinforcer value, with the stimulus associated with one pellet of food (2kHz) having a value of "1", and the stimulus associated with two pellets of food (4kHz) having a value of "2".

For the other contingency group, trained to associate responding to a 2kHz tone with delivery of two pellets of food (*2kHz=2pell*), the reference tonal stimuli (2kHz & 4kHz) were again assigned a numerical value equivalent to the number of pellets associated with that stimulus (so, 2kHz = "2", and 4kHz = "1"), with the values corresponding to the other (probe) stimuli appearing at intervals equivalent to those separating the probe stimulus values in the *4kHz=2pell* group, *with respect to the Hertz value of the reference stimuli.* So, the interval between "1" and the next probe value was greater when "1" represented 2kHz (as opposed to 4kHz). See Table 2.2 and Figure 2.10 for numerical and graphical representations of this scale, respectively.

APPENDIX C: ANALYSIS IN SPSS OF THE SINGLE-FREQUENCY PROBE SESSION DATA FROM CHAPTER 2

Lever choice

Method: fitting a probit function to model lever choice

For the analysis of *lever choice* we conducted in SPSS, we performed a probit analysis (M. Bateson, personal communication) on some of the data, which fits a cumulative normal distribution function, and then analysed aspects of this function in repeated-measures ANOVAs.

The term 'probit' was first coined by Bliss (1934) when developing statistical techniques to model and predict the amount of poison needed to kill insect pests, and this example illustrates the type of datasets probit analyses are often called upon to model: namely, the probability of a discrete event occurring (e.g. living as opposed to dying; pressing lever A as opposed to lever B; etc.) over increasing levels of a given stimulus (e.g. Finney, 1971; Norušis, 1999; Sokal & Rohlf, 1995); indeed, the technique has been used to analyse data from experimental paradigms similar to ours (Matheson et al., 2008). We used probit analyses to model the lever choice data from the nine probes intermediate to the two reference stimuli[177], since responding to these tones is more likely, *a priori*, to be cumulative: from scant responding at one end (i.e. to the probe stimulus neighbouring the reference tone to which responses on that lever are not reinforced), graduating to the highest frequency of responding at the other.

For each *subject,* the total number of presses on the lever associated with 2 pellets of food was taken for each of the *probe values* during each *measurement phase* (i.e. the count data were pooled for each *subject* across the three sessions in each *measurement phase*, with a maximum of 18 observations for each *probe value*, per *subject*, per *measurement phase*). These data were submitted to a probit analysis, with *probe value* (the standardised scale described on p.71) specified as the

[177] As opposed to *all* the *probe values* – i.e. including those on the far-side of the reference tones as well.

covariate[178], and an assumed natural response rate of zero (i.e. the analysis assumed a lever press would not be made in the absence of a probe stimulus).

We then submitted three different aspects of the resulting functions to repeated-measures general linear models (this procedure is somewhat similar to that employed by Matheson et al., 2008):

- the *probe value* at which the probit analysis estimated a 0.5 probability of the lever being pressed[179];

- the estimated probability of pressing the lever associated with 2 pellets of food for the intermediate-frequency *probe value* closest to the '2-pellet' reference tone;

- and the estimated probability of pressing the lever associated with 2 pellets of food for the intermediate-frequency *probe value* closest to the '1-pellet' reference tone.

See Figure 7.2 and Figure 7.3 for the hypothesised difference in these functions for *subjects* responding in an 'optimistic' and 'pessimistic' manner. Whilst submitting a number of aspects of the same function to separate analyses heightens the risk of Type II errors, through a multiplicity of significance-testing (Grafen & Hails, 2002), it nonetheless allows us to gain a more comprehensive picture of any changes in the form of that function, and indeed other experiments employing similar paradigms have found treatment-related differences only in certain values of the probes intermediate to the reference stimuli, including those neighbouring the reference stimuli (Burman et al., 2008b; Harding et al., 2004).

---

[178] As is usual in probit analyses (e.g. Norušis, 1999), the covariate was log-transformed prior to the analysis: i.e. the probe value scale has undergone an initial log-transformation to account for the psychophysical character of the stimuli, then a second log-transformation to map the distribution of responses in a manner the probit can better model.

[179] One could call this the point of probable *bisection* (e.g. Church & Deluty, 1977), or *bias* (Matheson et al., 2008).

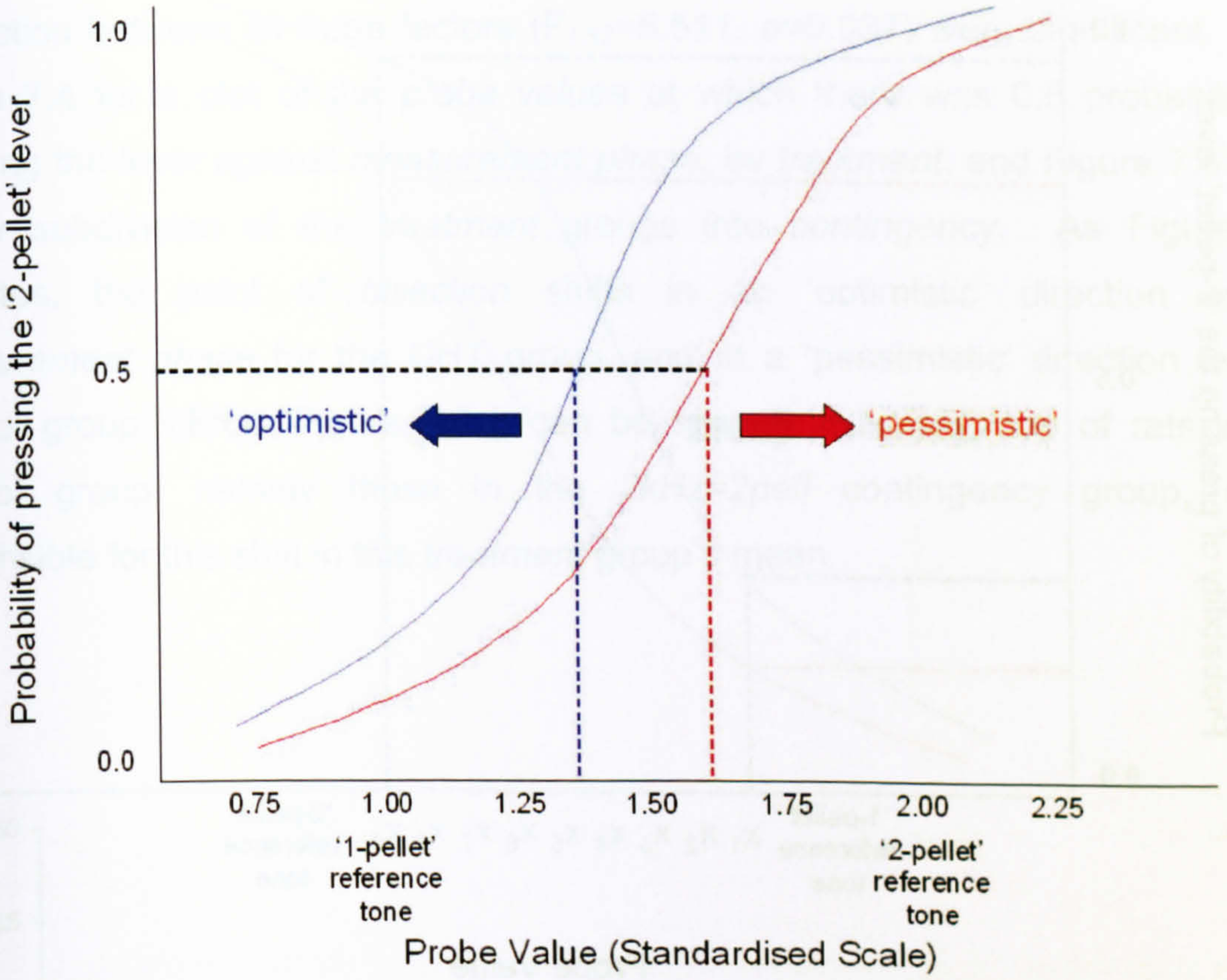

Figure 7.2 Schematic diagram illustrating the point of bisection (i.e. the *probe value* at which the probability of pressing one or other of the levers is 0.5) when responding in an 'optimistic' (blue curve), or 'pessimistic' (red curve), manner. Note, the *probe value* at the point of bisection is closer to the '1-pellet' reference tone when responding in an 'optimistic' manner than it is when responding in a 'pessimistic' manner.

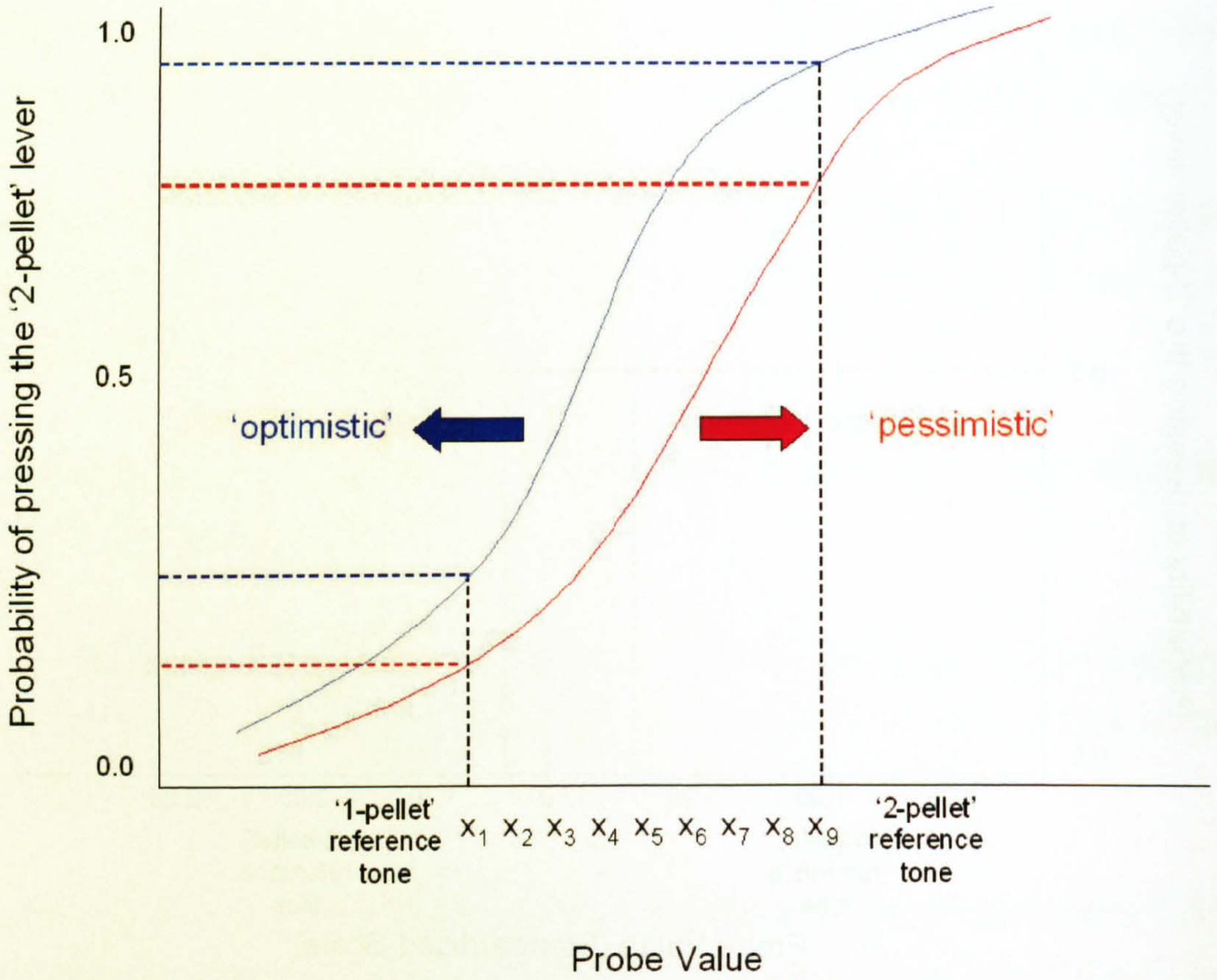

Figure 7.3 As Figure 7.2, but charting the hypothesised change in the probability of pressing the lever associated with 2 pellets of food for the *probe value* closest to the '1-pellet' reference tone ($\mathbf{x_1}$), and for the *probe value* closest to the '2-pellet' reference tone ($\mathbf{x_2}$).

Results: probit analysis & repeated measures anova

Firstly, the *probe value* at which the probit analysis estimated a 0.5 probability of the lever being pressed was compared in a repeated-measures GLM, with *measurement phase* as the within-subject factor, and *contingency* and *treatment* as the between-subject factors. This found no significant main effect of *measurement phase* ($F_{1,12}$=0.570, p=0.465) nor *treatment* ($F_{1,12}$=0.178, p=0.680), although *contingency* was significant ($F_{1,12}$=9.447, p=0.010), with the point of bisection for the *2kHz=2pell* group closer to the '2-pellet' reference tone (i.e. a bias towards 'pessimism': see Figure 7.2). The two-way interactions between *measurement phase* and *treatment* ($F_{1,12}$=2.835, p=0.118), and *contingency* and *treatment* ($F_{1,12}$=2.832, p=0.118) were not significant, whilst the two-way interaction between *measurement phase* and *contingency* ($F_{1,12}$=4.875, p=0.047), and the

interaction between all three factors ($F_{1,12}$=5.511, p=0.037) were significant. See Figure 7.4 for a plot of the *probe values* at which there was 0.5 probability of pressing the lever against *measurement phase*, by *treatment*, and Figure 7.5 for a further subdivision of the *treatment* groups into *contingency.* As Figure 7.4 indicates, the point of *bisection* shifts in an 'optimistic' direction across *measurement phase* for the UHT group, and in a 'pessimistic' direction for the Control group. From Figure 7.5 it can be seen that a subgroup of rats in the Control group, namely those in the *2kHz=2pell* contingency group, seem responsible for this shift in this *treatment* group's mean.

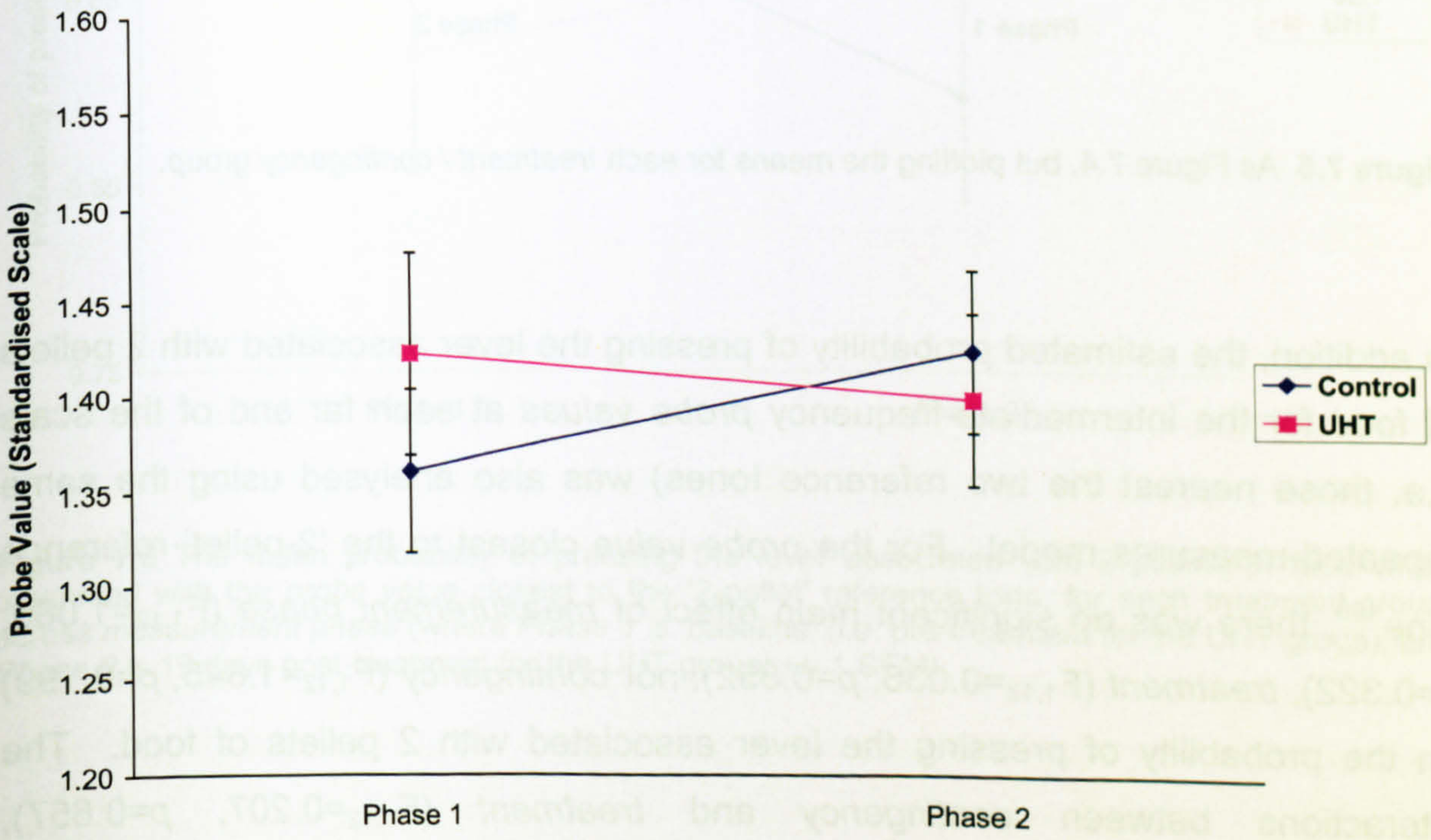

Figure 7.4 The mean probe value (a standardised scale in which '1' is the reference tone associated with 1 pellet of food, and '2' is the reference tone associated with 2 pellets of food) at which the probability of pressing the lever associated with 2 pellets of food was estimated as 0.5 by a probit analysis. The chart plots the means for each *treatment* group, across *measurement phase* (where *Phase 1* is 'baseline' (i.e. pre-treatment for the UHT group), and *Phase 2* is 19 days post-treatment for the UHT group; +/- 1 SEM). As Figure 7.2 illustrates, the closer to '1' this value is, the more 'optimistic' the style of responding.

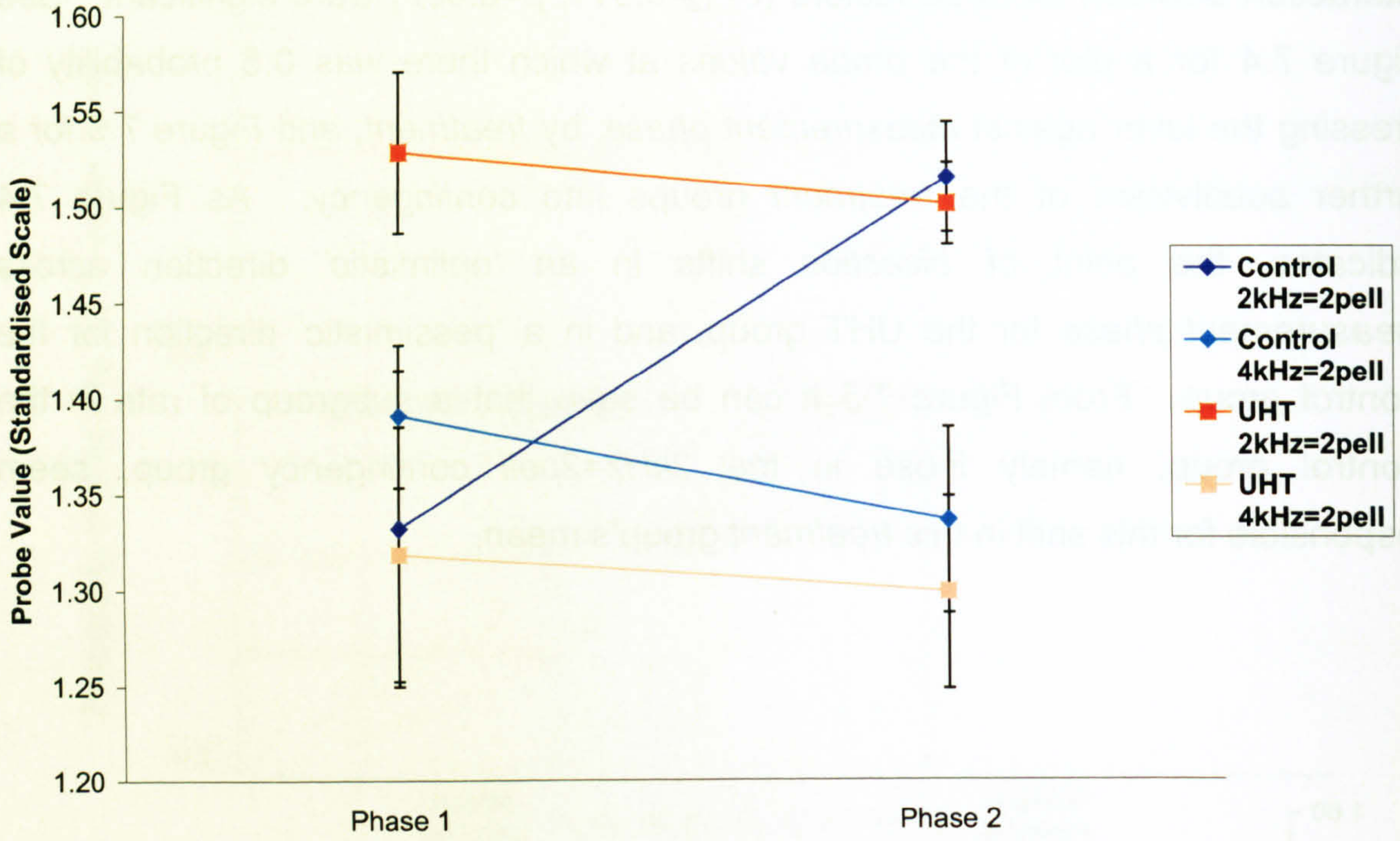

Figure 7.5 As Figure 7.4, but plotting the means for each *treatment / contingency* group.

In addition, the estimated probability of pressing the lever associated with 2 pellets of food for the intermediate-frequency *probe values* at each far end of the scale (i.e. those nearest the two reference tones) was also analysed using the same repeated-measures model. For the *probe value* closest to the '2-pellet' reference tone[180], there was no significant main effect of *measurement phase* ($F_{1,12}$=1.066, p=0.322), *treatment* ($F_{1,12}$=0.036, p=0.852), nor *contingency* ($F_{1,12}$=1.845, p=0.199) on the probability of pressing the lever associated with 2 pellets of food. The interactions between *contingency* and *treatment* ($F_{1,12}$=0.207, p=0.657), *measurement phase* and *contingency* ($F_{1,12}$=0.428, p=0.525), and all three factors ($F_{1,12}$=0.002, p=0.966) were not significant, whilst the two-way interaction between *measurement phase* and *treatment* ($F_{1,12}$=7.795, p=0.016) was significant. As Figure 7.6 indicates, the probability of pressing the lever associated with 2 pellets

[180] N.B. all assumptions of the GLM were met, bar the variances of the residuals from *Phase 2*, which failed a formal test of homogeneity (Levene's: p<0.001). Inspection of the residual plots indicated that the variance was particularly large for one of the *contingency / treatment* subgroups, which had middling predicted values. Homogeneity could not be achieved to the formal test's satisfaction using any of the transformations we explored. However, when the analysis was run again, without *contingency* as a factor, and all assumptions of the GLM satisfied, the significant *measurement phase*treatment* interaction remained ($F_{1,14}$=8.779, p=0.010): i.e. the significant finding quoted in the text appears to be robust.

of food when presented with the *probe value* closest to the '2-pellet' reference tone decreased (in a 'pessimistic' direction: see Figure 7.3) across *measurement phase* for the *Control* group, and increased (in an 'optimistic' direction) across *measurement phase* for the UHT group. Figure 7.7 plots this data further subdivided into *contingency*.

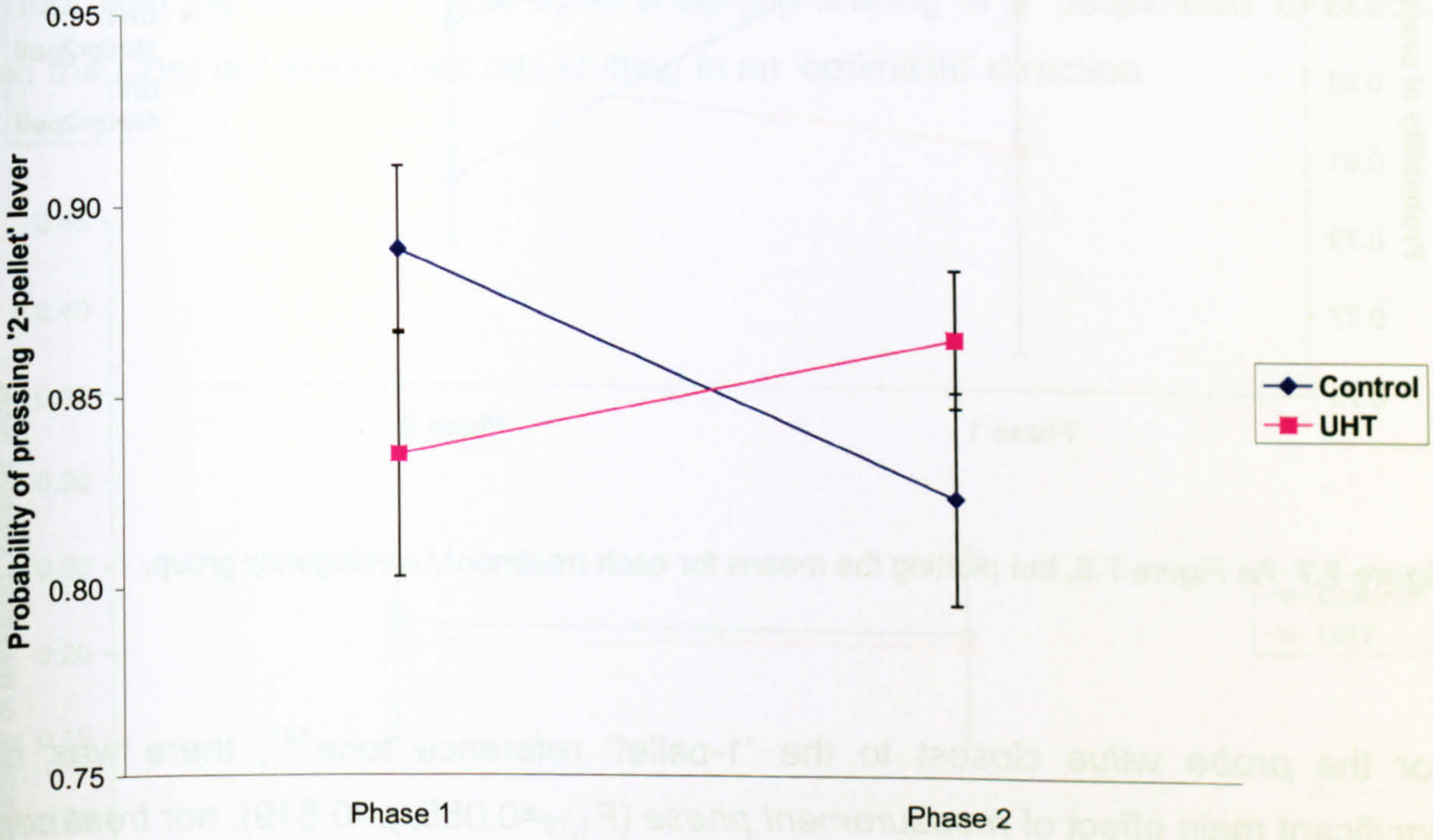

Figure 7.6 The mean probability of pressing the lever associated with 2 pellets of food when presented with the *probe value* closest to the '2-pellet' reference tone, for each *treatment* group, across *measurement phase* (where *Phase 1* is 'baseline' (i.e. pre-treatment for the UHT group), and *Phase 2* is 19 days post-treatment for the UHT group; +/- 1 SEM).

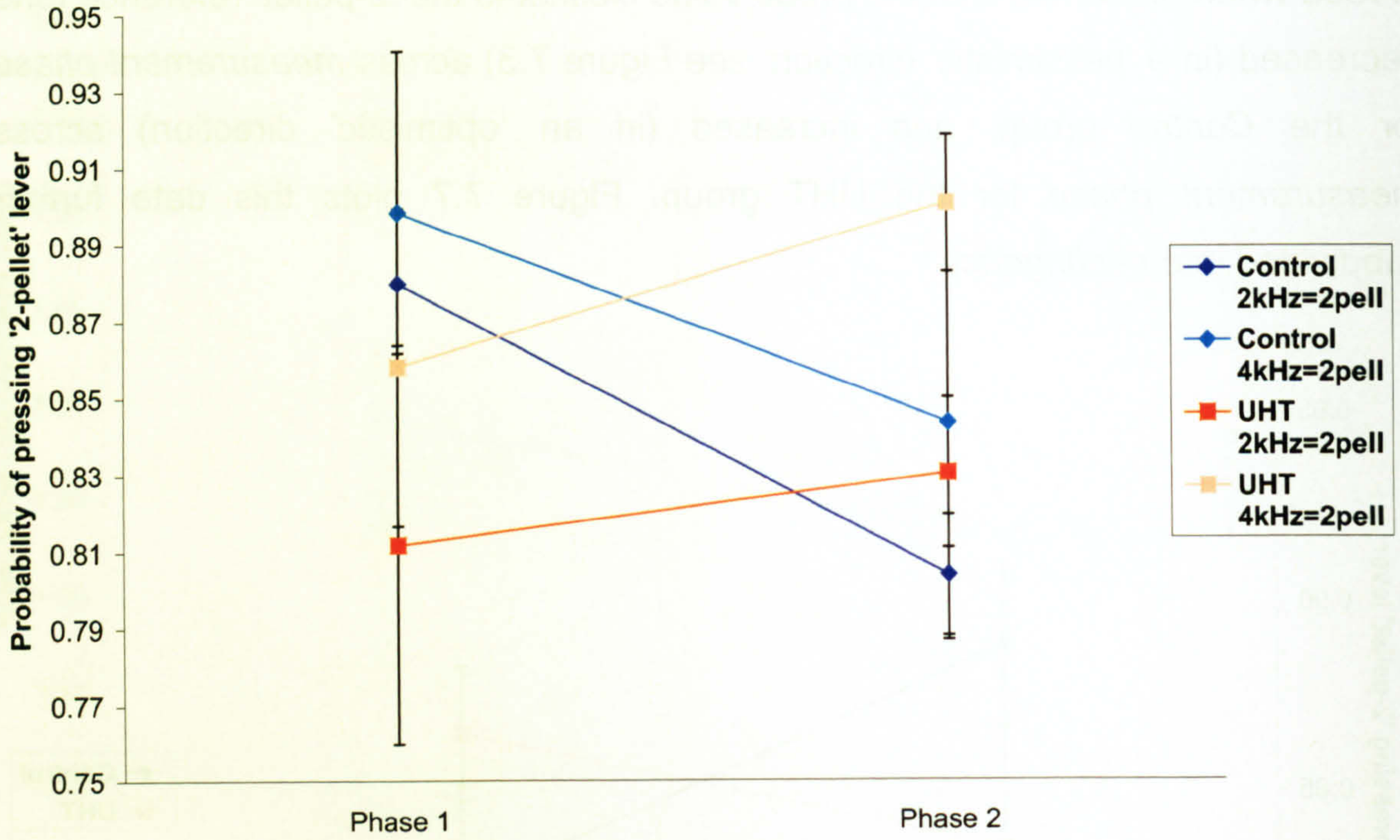

Figure 7.7 As Figure 7.6, but plotting the means for each *treatment / contingency* group.

For the *probe value* closest to the '1-pellet' reference tone[181], there was no significant main effect of *measurement phase* ($F_{1,12}$=0.055, p=0.819), nor *treatment* ($F_{1,12}$=0.094, p=0.764) on the (arc-sine-square-root-transformed) probability of pressing the lever associated with 2 pellets of food, although *contingency* ($F_{1,12}$=24.705, p<0.001) was highly-significant, with a lower probability of the *2kHz=2pell* pressing the lever associated with 2 pellet of food (i.e. a bias towards 'pessimism'). The interactions between *contingency* and *treatment* ($F_{1,12}$=2.465, p=0.142), and *measurement phase* and *treatment* ($F_{1,12}$=0.083, p=0.778), were not significant, whilst the interactions between *measurement phase* and *contingency*

[181] N.B. as before, all assumptions of the GLM were met, bar the variances of the residuals from *Phase 1*, which failed a formal test of homogeneity (Levene's: p<0.001). Again, inspection of the residual plots indicated that the variance was particularly large for the middling predicted values. Whilst homogeneity was not achieved to the formal test's satisfaction by the transformation we chose, the residual plots were, to our eyes, improved. Otherwise, when the analysis was run again, with *contingency* as the only between-subjects factor (and all assumptions of the GLM satisfied – we used a square-root transformation), the significant main effect of *contingency* remained ($F_{1,14}$=24.410, p<0.001), although the interaction between *contingency* and *measurement phase* was not significant ($F_{1,14}$=3.882, p=0.069). This indicates either that the heteroscedasticity present in the analysis described in the text has biased its outcome, and/or that the variance associated with *treatment* needs to be factored into the analysis for the interaction between *contingency* and *measurement phase* to be significant. There are grounds, then, to question the exact significance level of the three-way interaction between *measurement phase*, *treatment* and *contingency*, however, at p=0.008, this level is high, and Figure 7.9 suggests that this finding is reasonable, and rather likely.

($F_{1,12}$=6.259, p=0.028), and between all three factors ($F_{1,12}$=10.080, p=0.008) were significant. As Figure 7.8 indicates, the probability of pressing the lever associated with 2 pellets of food when presented with the *probe value* closest to the '1-pellet' reference tone changes little across *measurement phase* for each *treatment* group; however, as Figure 7.9 indicates, there is a marked difference across *measurement phase* for the different *contingency* subgroups within the *Control* group, with the *Control / 2kHz=2pell* subgroup shifting in a 'pessimistic' direction, and the *Control / 4kHz=2pell* rats shifting in an 'optimistic' direction.

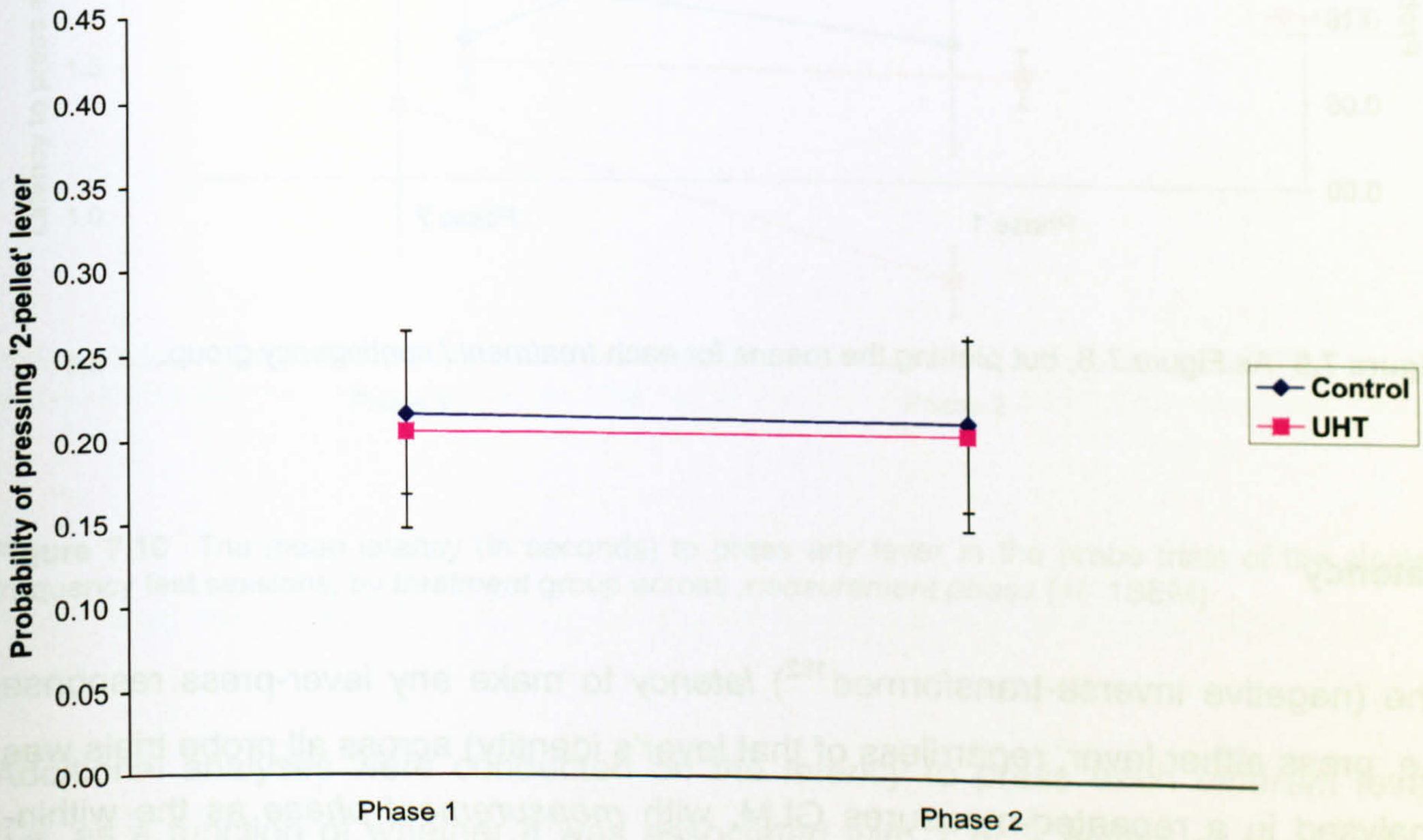

Figure 7.8 The mean probability of pressing the lever associated with 2 pellets of food when presented with the *probe value* closest to the '1-pellet' reference tone, for each *treatment* group, across *measurement phase* (+/-1SEM).

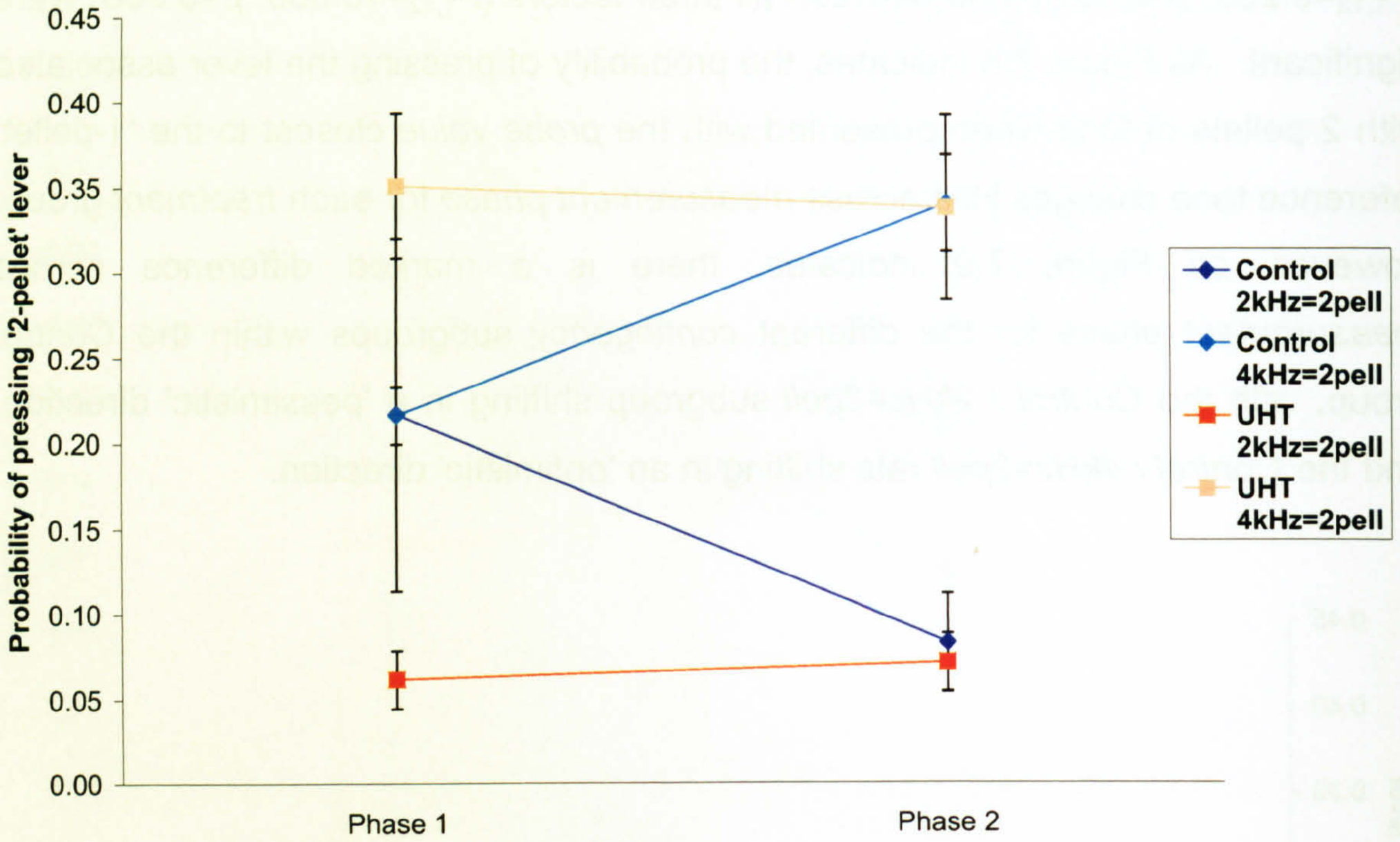

Figure 7.9 As Figure 7.8, but plotting the means for each *treatment / contingency* group.

Latency

The (negative inverse-transformed[182]) *latency* to make any lever-press response (i.e. press either lever, regardless of that lever's identity) across all probe trials was analysed in a repeated-measures GLM, with *measurement phase* as the within-subject factor, and *contingency* and *treatment* as the between-subject factors. This found no significant main effects nor interactions, although *treatment* neared significance at the 0.05 level (*measurement phase:* $F_{1,12}$=1.139, p=0.307; *treatment:* $F_{1,12}$=4.333, p=0.059; *contingency:* $F_{1,12}$<0.001, p=0.997; *measurement phase * treatment:* $F_{1,12}$=2.030, p=0.180; *measurement phase * contingency*: $F_{1,12}$=3.008, p=0.108; *treatment * contingency:* $F_{1,12}$=0.153, p=0.703; *measurement phase * contingency * treatment:* $F_{1,12}$=0.191, p=0.670). Figure 7.10, which plots the mean latency to press any lever in the probe trials of the single-frequency test

[182] i.e. -1/x (the minus sign preserves the order of the datapoints, which are otherwise reversed in an inverse transformation of 1/x) (e.g. Grafen & Hails, 2002).

sessions, illustrates why the main effect of *treatment* neared significance, with the Control group consistently slower to make a lever-press response.

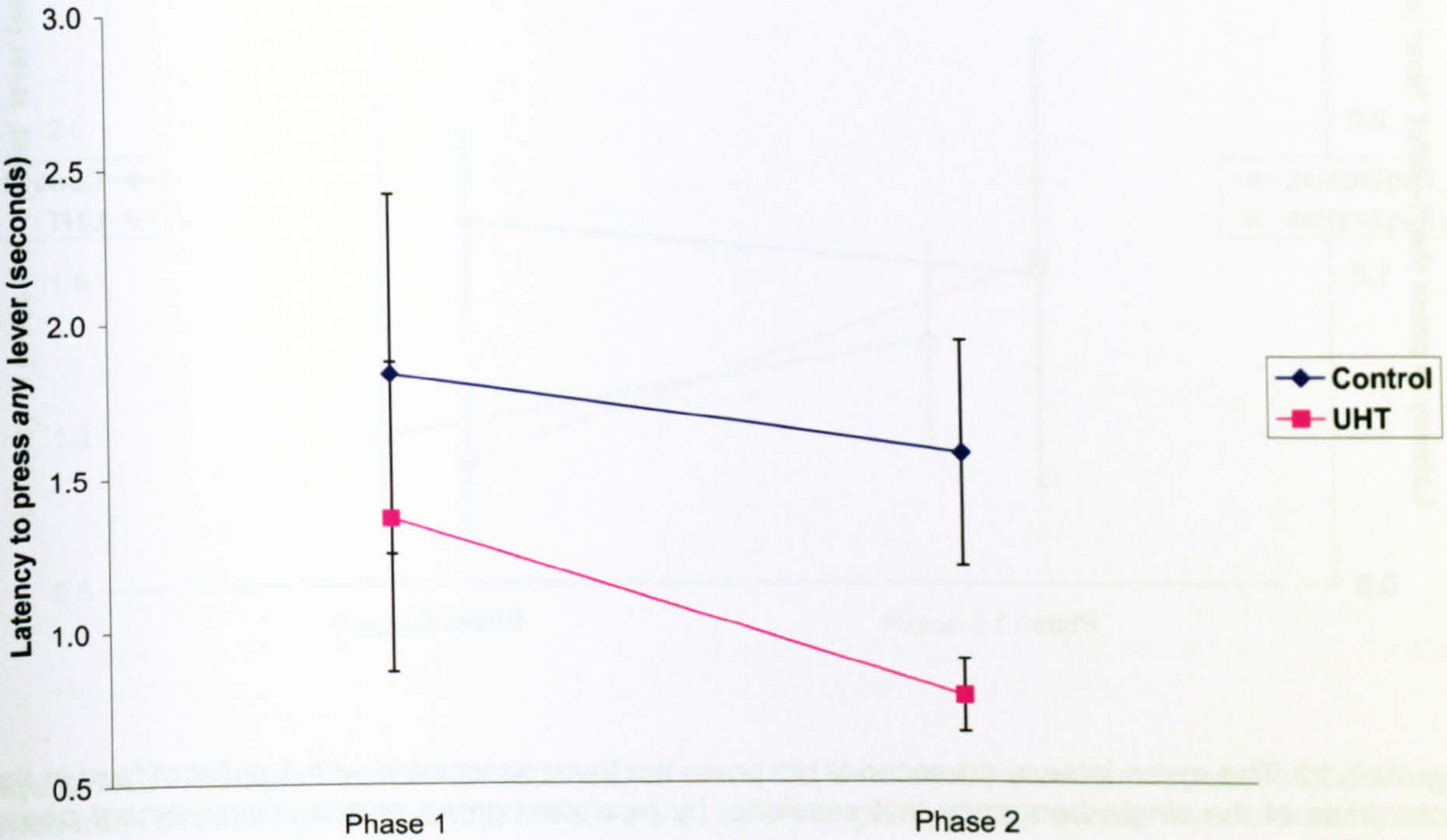

Figure 7.10 The mean latency (in seconds) to press *any* lever in the probe trials of the single-frequency test sessions, by *treatment* group across *measurement phase* (+/- 1SEM).

Additional analyses were conducted on the *latency* to press each *different* lever (i.e. as a function of whether it was associated with 1 or 2 pellets of food). The (log-transformed) *latency* to press the lever associated with 1 pellet of food across all single-frequency probe trials was analysed using the same repeated-measures model; this found no significant main effect nor interactions, although the two-way interaction between *measurement phase* and *contingency* neared significance at the 0.05 level (*measurement phase:* $F_{1,12}$=0.223, p=0.645; *treatment:* $F_{1,12}$=0.378, p=0.126; *contingency:* $F_{1,12}$=1.079, p=0.319; *measurement phase * treatment:* $F_{1,12}$=2.656, p=0.129; *measurement phase * contingency:* $F_{1,12}$=4.720, p=0.051; *treatment * contingency:* $F_{1,12}$=0.872, p=0.369; *measurement phase * contingency * treatment:* $F_{1,12}$=1.361, p=0.266). Figure 7.11 plots the mean latency to press the lever associated with 1 pellet of food for each of the *treatment* groups across *measurement phase,* whilst Figure 7.12 plots this data by *contingency* group.

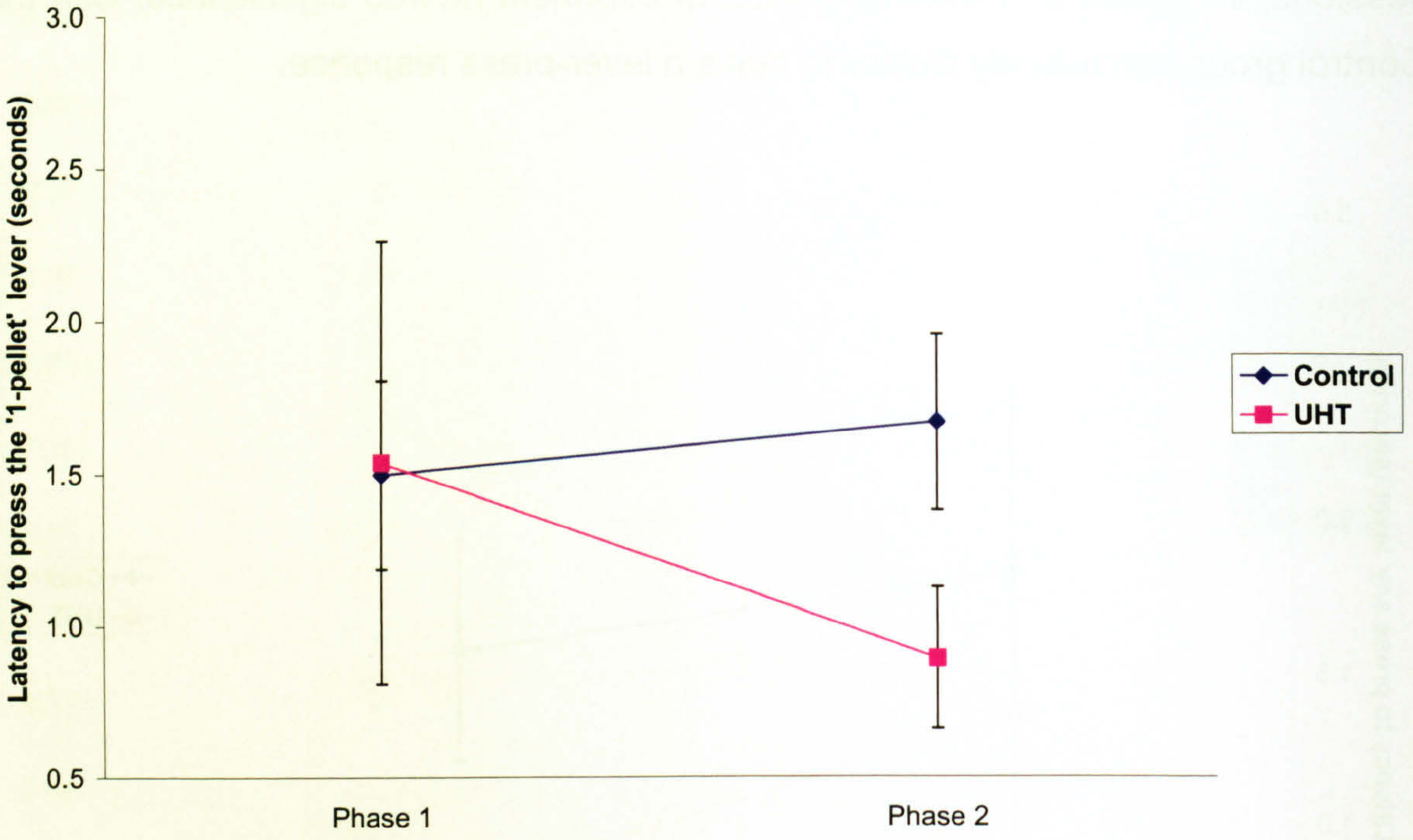

Figure 7.11 The mean latency (in seconds) to press the lever associated with 1 pellet of food in the probe trials of the single-frequency test sessions, by *treatment* group across *measurement phase* (+/- 1SEM).

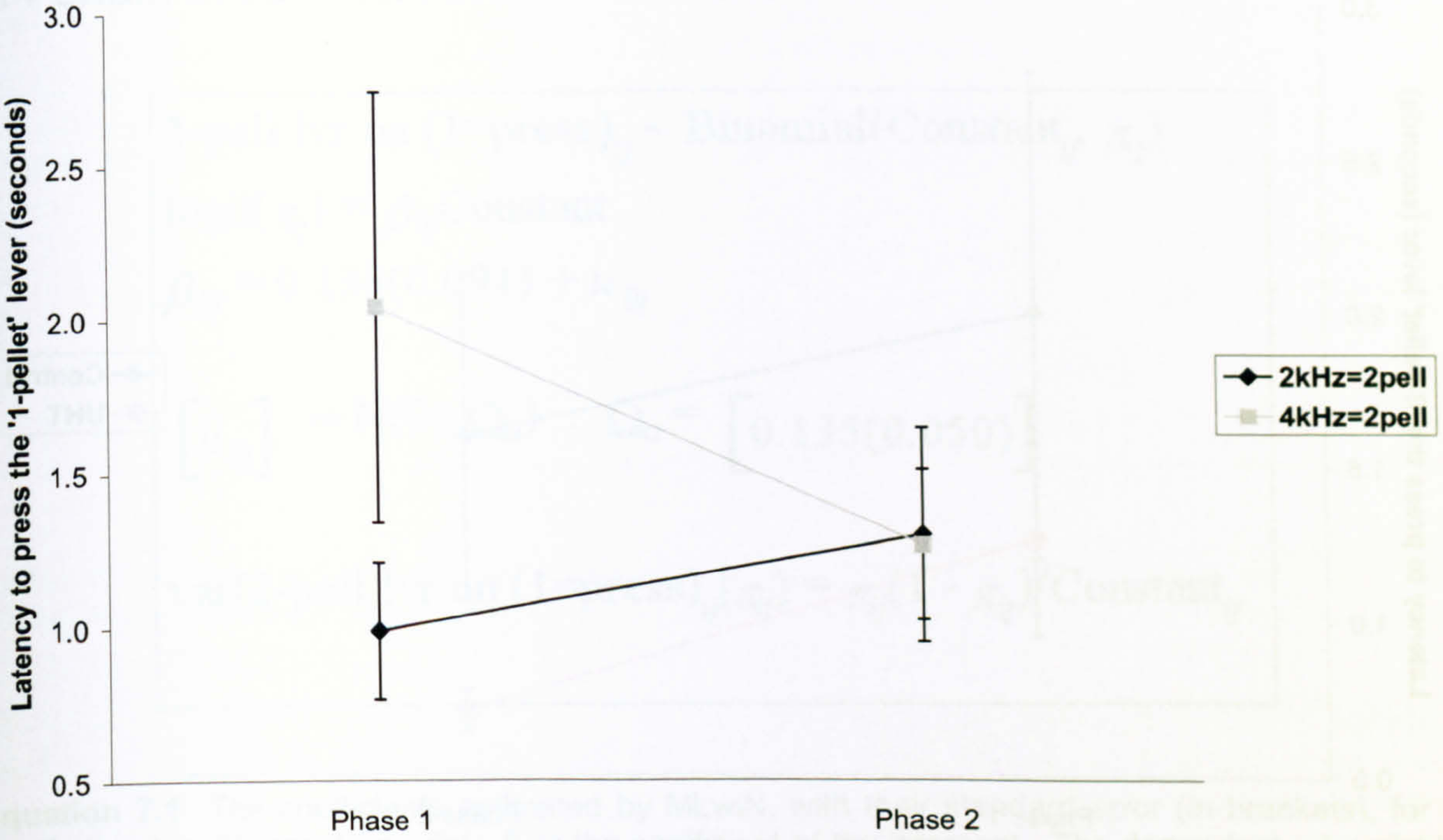

Figure 7.12 The mean latency (in seconds) to press the lever associated with 1 pellet of food in the probe trials of the single-frequency test sessions, by *contingency* group across *measurement phase* (+/- 1SEM).

The (negative inverse-transformed) *latency* to press the lever associated with 2 pellets of food across all probe trials was also analysed using the same repeated-measures model. Again, this found no significant main effects nor interactions (*measurement phase:* $F_{1,12}$=2.996, p=0.109; *treatment:* $F_{1,12}$=3.566, p=0.083; *contingency:* $F_{1,12}$=1.285, p=0.279; *measurement phase * treatment:* $F_{1,12}$=1.658, p=0.222; *measurement phase * contingency*: $F_{1,12}$=2.803, p=0.120; *treatment * contingency:* $F_{1,12}$=0.003, p=0.959; *measurement phase * contingency * treatment:* $F_{1,12}$=0.022, p=0.883). Figure 7.13 plots the mean latency to press the lever associated with 2 pellets of food for each *treatment* group across *measurement phase.*

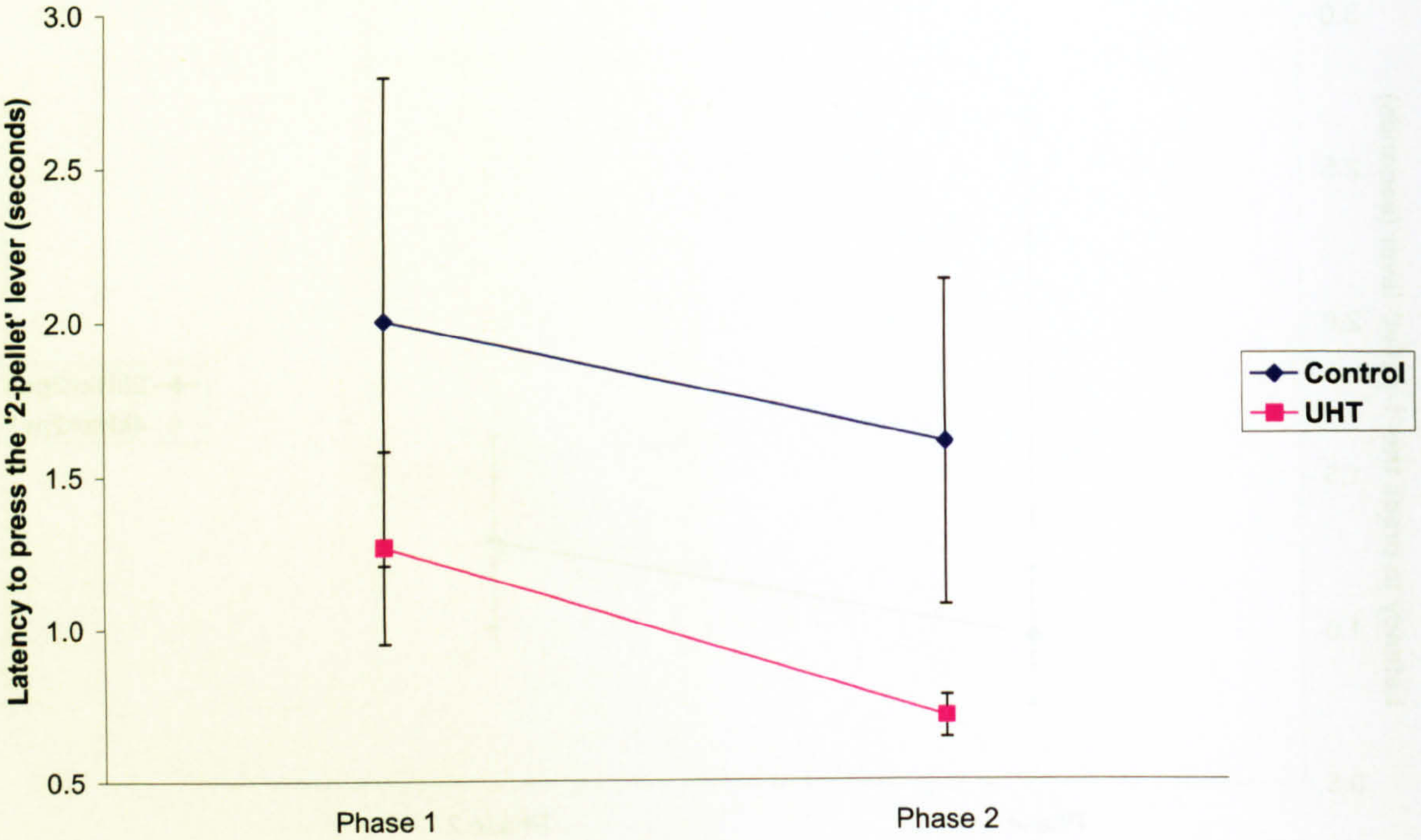

Figure 7.13 The mean latency (in seconds) to press the lever associated with 2 pellet of food in the probe trials of the single-frequency test sessions, by *treatment* group across *measurement phase* (+/- 1SEM).

APPENDIX D: SUPPLEMENTARY RESULTS

$$\text{2-pell lvr on (1=press)}_{ij} \sim \text{Binomial}(\text{Constant}_{ij}, \pi_{ij})$$

$$\text{logit}(\pi_{ij}) = \beta_{0j}\text{Constant}$$

$$\beta_{0j} = 0.138(0.094) + u_{0j}$$

$$\begin{bmatrix} u_{0j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0.135(0.050) \end{bmatrix}$$

$$\text{var(2-pell lvr on (1=press)}_{ij} | \pi_{ij}) = \pi_{ij}(1 - \pi_{ij})/\text{Constant}_{ij}$$

Equation 7.1 The coefficients estimated by MLwiN, with their standard error (in brackets), for a simple 'random intercepts' model. β_0 is the coefficient of the intercept. The dependent (y) variable is a press (=1) or not (=0) on the lever associated with 2 pellets of food.

$$\text{2-pell lvr on (1=press)}_{ij} \sim \text{Binomial}(\text{Constant}_{ij}, \pi_{ij})$$

$$\text{logit}(\pi_{ij}) = \beta_{0j}\text{Constant} + \beta_{1j}\text{Standard log scale}_{ij}$$

$$\beta_{0j} = 0.174(0.105) + u_{0j}$$

$$\beta_{1j} = 2.766(0.193) + u_{1j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0.164(0.062) & \\ -0.144(0.088) & 0.524(0.211) \end{bmatrix}$$

$$\text{var(2-pell lvr on (1=press)}_{ij} | \pi_{ij}) = \pi_{ij}(1 - \pi_{ij})/\text{Constant}_{ij}$$

Equation 7.2 The coefficients estimated by MLwiN, with their standard error (in brackets), for a simple 'random slope' model. B_1 is the coefficient of the added term, which is the standardised log scale for *probe value*.

2-pell lvr on (1=press)$_{ij}$ ~ Binomial(Constant$_{ij}$, π_{ij})

logit(π_{ij}) = β_{0j}Constant + β_{1j}Standard log scale$_{ij}$ + β_{2j}Standard log scale quadratic$_{ij}$ + -5.993(0.368)Standard log scale cubic$_{ij}$

$\beta_{0j} = 0.341(0.128) + u_{0j}$

$\beta_{1j} = 4.858(0.233) + u_{1j}$

$\beta_{2j} = -0.834(0.425) + u_{2j}$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) \; : \; \Omega_u = \begin{bmatrix} 0.242(0.093) & & \\ -0.299(0.124) & 0.522(0.212) & \\ -0.566(0.268) & 0.869(0.395) & 2.558(1.024) \end{bmatrix}$$

var(2-pell lvr on (1=press)$_{ij}$| π_{ij}) = $\pi_{ij}(1 - \pi_{ij})$/Constant$_{ij}$

Equation 7.3 The coefficient estimates generated by MLwiN for the specified model. Here, the 'slope' terms (β_1 β_2 & β_3) are the probe stimuli values, standardised around the quantity of food reinforcer, fitted up to a cubic power.

2-pell lvr on (1=press)$_{ij}$ ~ Binomial(Constant$_{ij}$, π_{ij})

logit(π_{ij}) = β_{0j}Constant + β_{1j}Standard log scale$_{ij}$ + β_{2j}Standard log scale quadratic$_{ij}$ + -5.922(0.368)Standard log scale cubic$_{ij}$ + 0.467(0.137)4kHz=2pell$_j$

$\beta_{0j} = 0.103(0.116) + u_{0j}$

$\beta_{1j} = 4.836(0.235) + u_{1j}$

$\beta_{2j} = -0.802(0.428) + u_{2j}$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) \; : \; \Omega_u = \begin{bmatrix} 0.114(0.048) & & \\ -0.193(0.086) & 0.534(0.216) & \\ -0.268(0.176) & 0.898(0.403) & 2.583(1.033) \end{bmatrix}$$

var(2-pell lvr on (1=press)$_{ij}$| π_{ij}) = $\pi_{ij}(1 - \pi_{ij})$/Constant$_{ij}$

Equation 7.4 The coefficient estimates generated by MLwiN for the specified model (modelling data from both *measurement phases*).

2-pell lvr on (1=press)$_{ij}$ ~ Binomial(Constant$_{ij}$, π_{ij})

logit(π_{ij}) = β_{0j}Constant + β_{1j}Standard log scale$_{ij}$ + β_{2j}Standard log scale quadratic$_{ij}$ + -7.971(0.565)Standard log scale cubic$_{ij}$ + 0.765(0.168)4kHz=2pell$_{j}$ + -1.956(0.412)4kHz=2pell.Standard log scale$_{ij}$ + -2.680(0.549)4kHz=2pell.Standard log scale quadratic$_{ij}$ + 3.204(0.768)4kHz=2pell.Standard log scale cubic$_{ij}$

β_{0j} = -0.076(0.119) + u_{0j}

β_{1j} = 5.928(0.301) + u_{1j}

β_{2j} = 0.906(0.391) + u_{2j}

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) \quad : \quad \Omega_u = \begin{bmatrix} 0.091(0.040) & & \\ -0.124(0.060) & 0.307(0.133) & \\ -0.065(0.091) & 0.365(0.182) & 0.731(0.374) \end{bmatrix}$$

var(2-pell lvr on (1=press)$_{ij}$|π_{ij}) = $\pi_{ij}(1 - \pi_{ij})$/Constant$_{ij}$

Equation 7.5 The coefficient estimates generated by MLwiN for the specified model (modelling data from both *measurement phases*).

2-pell lvr on (1=press)$_{ij}$ ~ Binomial(Constant$_{ij}$, π_{ij})

logit(π_{ij}) = β_{0j}Constant + β_{1j}Standard log scale$_{ij}$ + β_{2j}Standard log scale quadratic$_{ij}$ + -7.981(0.565)Standard log scale cubic$_{ij}$ + 0.766(0.169)4kHz=2pell$_{j}$ + -1.958(0.412)4kHz=2pell.Standard log scale$_{ij}$ + -2.684(0.548)4kHz=2pell.Standard log scale quadratic$_{ij}$ + 3.208(0.768)4kHz=2pell.Standard log scale cubic$_{ij}$ + -0.142(0.052)Phase2$_{ij}$

β_{0j} = -0.005(0.122) + u_{0j}

β_{1j} = 5.933(0.302) + u_{1j}

β_{2j} = 0.908(0.390) + u_{2j}

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) \quad : \quad \Omega_u = \begin{bmatrix} 0.091(0.040) & & \\ -0.125(0.060) & 0.307(0.132) & \\ -0.065(0.091) & 0.368(0.182) & 0.727(0.370) \end{bmatrix}$$

var(2-pell lvr on (1=press)$_{ij}$|π_{ij}) = $\pi_{ij}(1 - \pi_{ij})$/Constant$_{ij}$

Equation 7.6 The coefficient estimates generated by MLwiN for the specified model (modelling data from both *measurement phases*).

2-pell lvr on (1=press)$_{ij}$ ~ Binomial(Constant$_{ij}$, π_{ij})

logit(π_{ij}) = β_{0j}Constant + β_{1j}Standard log scale$_{ij}$ + β_{2j}Standard log scale quadratic$_{ij}$ + -8.020(0.567)Standard log scale cubic$_{ij}$ + 0.568(0.177)4kHz=2pell$_{j}$ + -1.992(0.414)4kHz=2pell.Standard log scale$_{ij}$ + -2.687(0.550)4kHz=2pell.Standard log scale quadratic$_{ij}$ + 3.253(0.769)4kHz=2pell.Standard log scale cubic$_{ij}$ + -0.351(0.077)Phase2$_{ij}$ + 0.395(0.105)Phase2.4kHz=2pell$_{ij}$

β_{0j} = 0.099(0.126) + u_{0j}

β_{1j} = 5.964(0.303) + u_{1j}

β_{2j} = 0.913(0.392) + u_{2j}

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) \quad : \quad \Omega_u = \begin{bmatrix} 0.092(0.040) & & \\ -0.126(0.060) & 0.310(0.134) & \\ -0.066(0.092) & 0.371(0.184) & 0.735(0.374) \end{bmatrix}$$

var(2-pell lvr on (1=press)$_{ij}$|π_{ij}) = $\pi_{ij}(1 - \pi_{ij})$/Constant$_{ij}$

Equation 7.7 The coefficient estimates generated by MLwiN for the specified model (modelling data from both *measurement phases*).

2-pell lvr on (1=press)$_{ij}$ ~ Binomial(Constant$_{ij}$, π_{ij})

logit(π_{ij}) = β_{0j}Constant + β_{1j}Standard log scale$_{ij}$ + β_{2j}Standard log scale quadratic$_{ij}$ + -8.028(0.567)Standard log scale cubic$_{ij}$ + 0.568(0.176)4kHz=2pell$_j$ +
-1.995(0.413)4kHz=2pell.Standard log scale$_{ij}$ + -2.688(0.550)4kHz=2pell.Standard log scale quadratic$_{ij}$ +
3.265(0.770)4kHz=2pell.Standard log scale cubic$_{ij}$ + -0.352(0.077)Phase2$_{ij}$ + 0.395(0.105)Phase2.4kHz=2pell$_{ij}$ + -0.062(0.133)UHT$_j$

β_{0j} = 0.130(0.142) + u_{0j}

β_{1j} = 5.968(0.303) + u_{1j}

β_{2j} = 0.914(0.392) + u_{2j}

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.090(0.039) & & \\ -0.124(0.060) & 0.309(0.133) & \\ -0.063(0.091) & 0.373(0.184) & 0.735(0.373) \end{bmatrix}$$

var(2-pell lvr on (1=press)$_{ij}$| π_{ij}) = $\pi_{ij}(1 - \pi_{ij})$/Constant$_{ij}$

Equation 7.8 The coefficient estimates generated by MLwiN for the specified model (modelling data from both *measurement phases*).

2-pell lvr on (1=press)$_{ij}$ ~ Binomial(Constant$_{ij}$, π_{ij})

logit(π_{ij}) = β_{0j}Constant + β_{1j}Standard log scale$_{ij}$ + β_{2j}Standard log scale quadratic$_{ij}$ + -8.073(0.570)Standard log scale cubic$_{ij}$ + 0.306(0.222)4kHz=2pell$_j$ +
-2.027(0.414)4kHz=2pell.Standard log scale$_{ij}$ + -2.684(0.551)4kHz=2pell.Standard log scale quadratic$_{ij}$ +
3.313(0.771)4kHz=2pell.Standard log scale cubic$_{ij}$ + -0.673(0.107)Phase2$_{ij}$ + 0.704(0.148)Phase2.4kHz=2pell$_{ij}$ + -0.525(0.211)UHT$_j$ +
0.536(0.286)UHT.4kHz=2pell$_j$ + 0.673(0.155)UHT.Phase2$_{ij}$ + -0.648(0.211)UHT.4kHz=2pell.Phase2$_{ij}$

β_{0j} = 0.356(0.160) + u_{0j}

β_{1j} = 5.999(0.304) + u_{1j}

β_{2j} = 0.913(0.393) + u_{2j}

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.081(0.036) & & \\ -0.111(0.056) & 0.308(0.134) & \\ -0.047(0.086) & 0.369(0.184) & 0.737(0.375) \end{bmatrix}$$

var(2-pell lvr on (1=press)$_{ij}$| π_{ij}) = $\pi_{ij}(1 - \pi_{ij})$/Constant$_{ij}$

Equation 7.9 The coefficient estimates generated by MLwiN for the specified model (modelling data from both *measurement phases*).

Ordered Subgroup	Observed (O)	Expected (E) (i.e. predicted from model)	$(O-E)^2$	$(O-E)^2 / E$
1	104	94.59	88.61	0.94
2	154	145.39	74.07	0.51
3	199	217.68	348.89	1.60
4	300	302.85	8.15	0.03
5	399	422.35	544.99	1.29
6	590	569.62	415.16	0.73
7	654	653.21	0.62	0.00
8	698	694.85	9.90	0.01
9	726	725.77	0.05	0.00
10	776	775.43	0.32	0.00
			χ^2 statistic:	**5.11**
			d.f.	**10**
			p	**0.884**

Table 7.2 Calculations pertaining to the Hosmer-Lemeshow goodness-of-fit test of the model specified in Equation 7.9; the 'Observed' and 'Expected' refer to the number of presses on the lever associated with 2 pellets of food, and the subgroups are ordered with respect to the 'Expected' values (with subgroup 1 having the lowest expected values, and subgroup 10 the greatest, with a roughly equal number of *trials* in each subgroup).

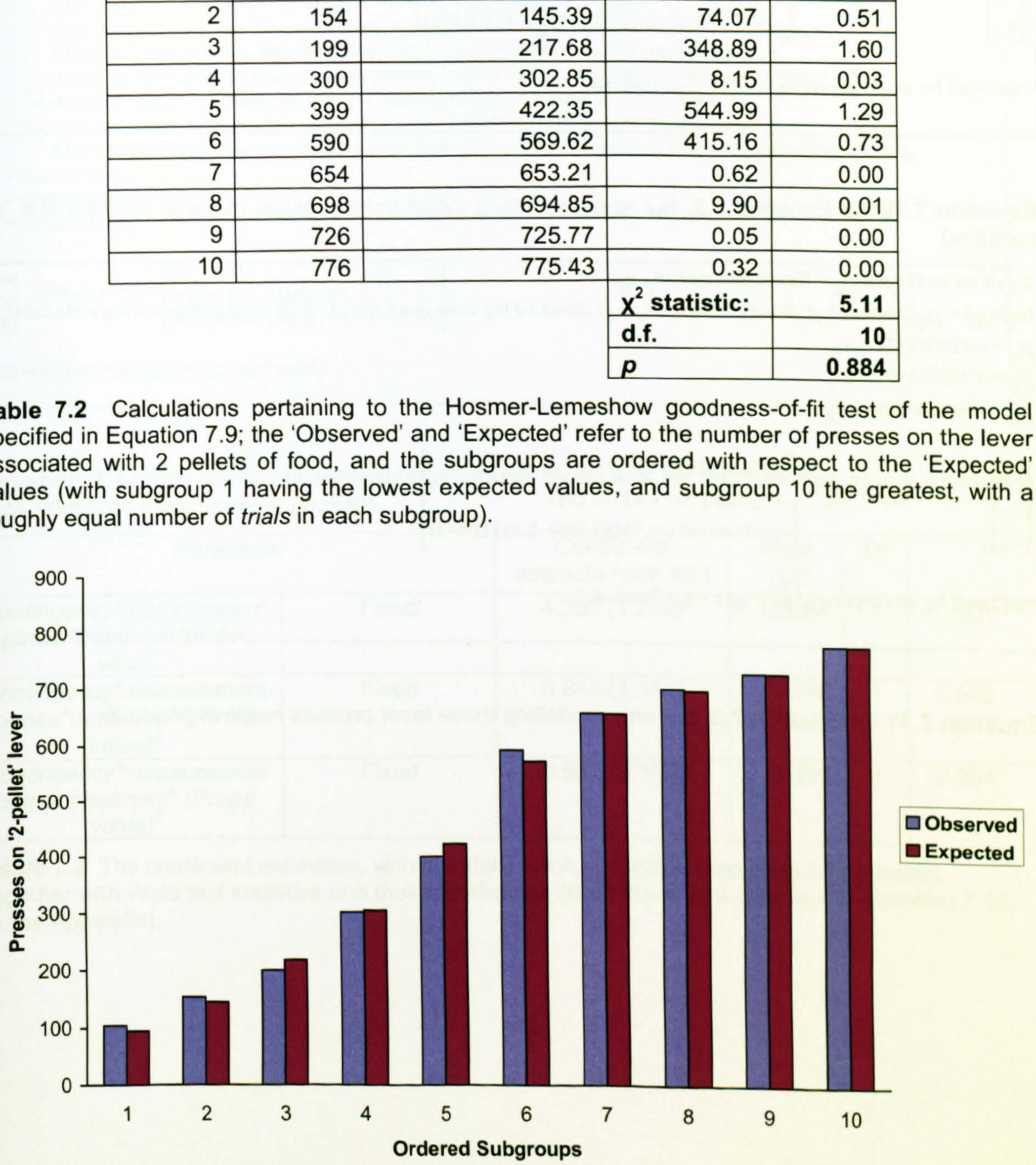

Figure 7.14 Plot of the 'Observed' and 'Expected' number of presses on the lever associated with 2 pellets of food, across the ordered subgroups (see Table 7.2).

2-pell lvr on (1=press)$_{ij}$ ~ Binomial(Constant$_{ij}$, π_{ij})

logit(π_{ij}) = β_{0j}Constant + β_{1j}Standard log scale$_{ij}$ + β_{2j}Standard log scale quadratic$_{ij}$ + -6.063(0 515)Standard log scale cubic$_{ij}$

β_{0j} = 0.368(0.137) + u_{0j}

β_{1j} = 4.959(0.289) + u_{1j}

β_{2j} = -0 580(0.454) + u_{2j}

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0\ 258(0.106) & & \\ -0.324(0.145) & 0.621(0.273) & \\ -0.487(0.290) & 0.987(0.470) & 2.611(1.157) \end{bmatrix}$$

var(2-pell lvr on (1=press)$_{ij}$| π_{ij}) = $\pi_{ij}(1 - \pi_{ij})$/Constant$_{ij}$

Equation 7.10 As Equation 7.3, but only modelling those lever presses made in *phase 1* (i.e. at baseline).

2-pell lvr on (1=press)$_{ij}$ ~ Binomial(Constant$_{ij}$, π_{ij})

logit(π_{ij}) = β_{0j}Constant + β_{1j}Standard log scale$_{ij}$ + β_{2j}Standard log scale quadratic$_{ij}$ + -5.797(0 488)Standard log scale cubic$_{ij}$

β_{0j} = 0.319(0.145) + u_{0j}

β_{1j} = 4.769(0.259) + u_{1j}

β_{2j} = -1.054(0.432) + u_{2j}

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0\ 294(0.119) & & \\ -0.270(0.128) & 0.421(0.198) & \\ -0.635(0.312) & 0.912(0.395) & 2.341(1.052) \end{bmatrix}$$

var(2-pell lvr on (1=press)$_{ij}$| π_{ij}) = $\pi_{ij}(1 - \pi_{ij})$/Constant$_{ij}$

Equation 7.11 As Equation 7.3, but only modelling those lever presses made in *phase 2*.

2-pell lvr on (1=press)$_{ij}$ ~ Binomial(Constant$_{ij}$, π_{ij})

logit(π_{ij}) = β_{0j}Constant + β_{1j}Standard log scale$_{ij}$ + β_{2j}Standard log scale quadratic$_{ij}$ + -6 409(1 112)Standard log scale cubic$_{ij}$ + 0 154(0.248)4kHz=2pell$_{ij}$ +
0 030(0 721)4kHz=2pell Standard log scale$_{ij}$ + -1.908(0.878)4kHz=2pell Standard log scale quadratic$_{ij}$ +
-0 481(1 555)4kHz=2pell Standard log scale cubic$_{ij}$ + -0 705(0 151)Phase2$_{ij}$ + 0 745(0 213)Phase2.4kHz=2pell$_{ij}$ + -0 809(0 249)UHT$_{ij}$ +
0 867(0 351)UHT 4kHz=2pell$_{ij}$ + 0 902(0 216)UHT Phase2$_{ij}$ + -0.742(0 303)UHT.4kHz=2pell Phase2$_{ij}$ + 0 682(0 630)Phase2 Standard log scale$_{ij}$ +
0 520(0 680)Phase2.Standard log scale quadratic$_{ij}$ + -1 899(1 561)Phase2.Standard log scale cubic$_{ij}$ + 1 933(0 767)UHT.Standard log scale$_{ij}$ +
2 227(0 891)UHT.Standard log scale quadratic$_{ij}$ + -3 540(1 652)UHT.Standard log scale cubic$_{ij}$ +
-2 399(0 858)4kHz=2pell Phase2 Standard log scale$_{ij}$ + -0.231(0 962)4kHz=2pell Phase2 Standard log scale quadratic$_{ij}$ +
5 455(2 168)4kHz=2pell Phase2 Standard log scale cubic$_{ij}$ + -3 919(1.023)4kHz=2pell UHT Standard log scale$_{ij}$ +
-1 871(1.247)4kHz=2pell UHT Standard log scale quadratic$_{ij}$ + 6 674(2 213)4kHz=2pell UHT.Standard log scale cubic$_{ij}$ +
-1 656(0.927)Phase2 UHT.Standard log scale$_{ij}$ + -1 983(0 991)Phase2 UHT Standard log scale quadratic$_{ij}$ +
3 938(2 293)Phase2 UHT Standard log scale cubic$_{ij}$ + 4 360(1 230)4kHz=2pell Phase2 UHT Standard log scale$_{ij}$ +
0 843(1 374)4kHz=2pell Phase2 UHT Standard log scale quadratic$_{ij}$ + -8 932(3 102)4kHz=2pell Phase2 UHT.Standard log scale cubic$_{ij}$

β_{0j} = 0 449(0.175) + u_{0j}
β_{1j} = 5 160(0 520) + u_{1j}
β_{2j} = 0.081(0.612) + u_{2j}

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0\ 077(0\ 035) & & \\ -0\ 095(0\ 051) & 0\ 253(0.116) & \\ -0\ 023(0.079) & 0\ 308(0.159) & 0\ 617(0\ 333) \end{bmatrix}$$

var(2-pell lvr on (1=press)$_{ij}$| π_{ij}) = $\pi_{ij}(1 - \pi_{ij})$/Constant$_{ij}$

Equation 7.12 The coefficient estimates generated by MLwiN for the specified model (modelling data from both *measurement phases*).

Parameter		Coefficient estimate (with SE)	Wald (x2)	Df	*P*
*Contingency*measurement phase*treatment*probe value*	Fixed	4.360 (1.230)	12.564	1	<0.001 **
*Contingency*measurement phase*treatment* (Probe value)^2*	Fixed	0.843 (1.374)	0.376	1	0.540
*Contingency*measurement phase*treatment* (Probe value)^3*	Fixed	-8.932 (3.102)	8.290	1	0.004

Table 7.3 The coefficient estimates, with standard error, of certain fixed parts of the model, together with Wald test statistics and their significance (from the model specified in Equation 7.12, in the Appendix).

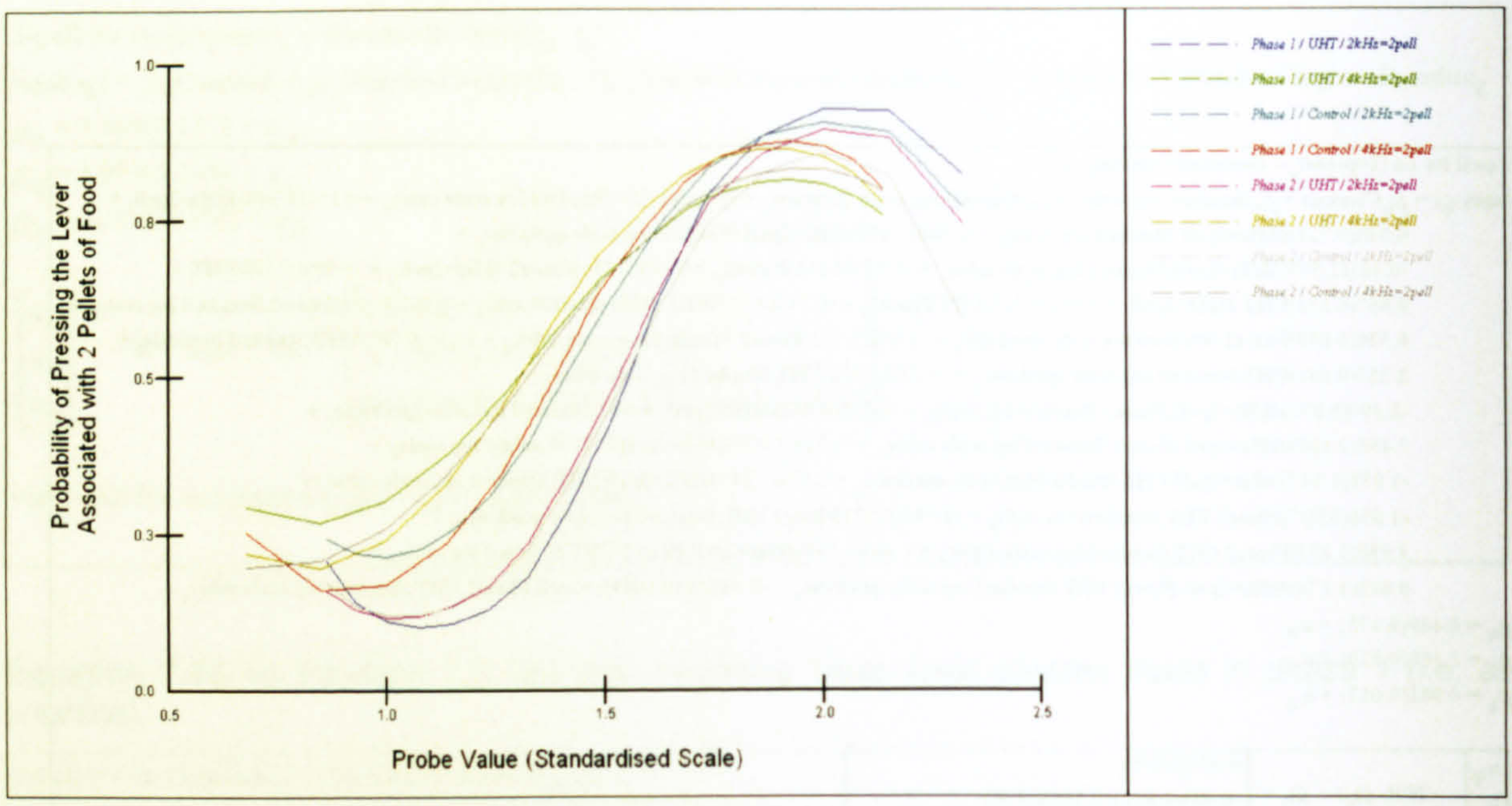

Figure 7.15 The probability of pressing the lever associated with 2 pellets of food (as opposed to the lever associated with 1 pellet of food) across different values of the probe stimuli (a standardised scale, with *1.0* corresponding to the reference tone associated with 1 pellet of food, and *2.0* corresponding to the reference tone associated with 2 pellets of food), by *measurement phase / treatment / contingency* group. These predicted lines were generated from the model and estimates in Equation 7.12.

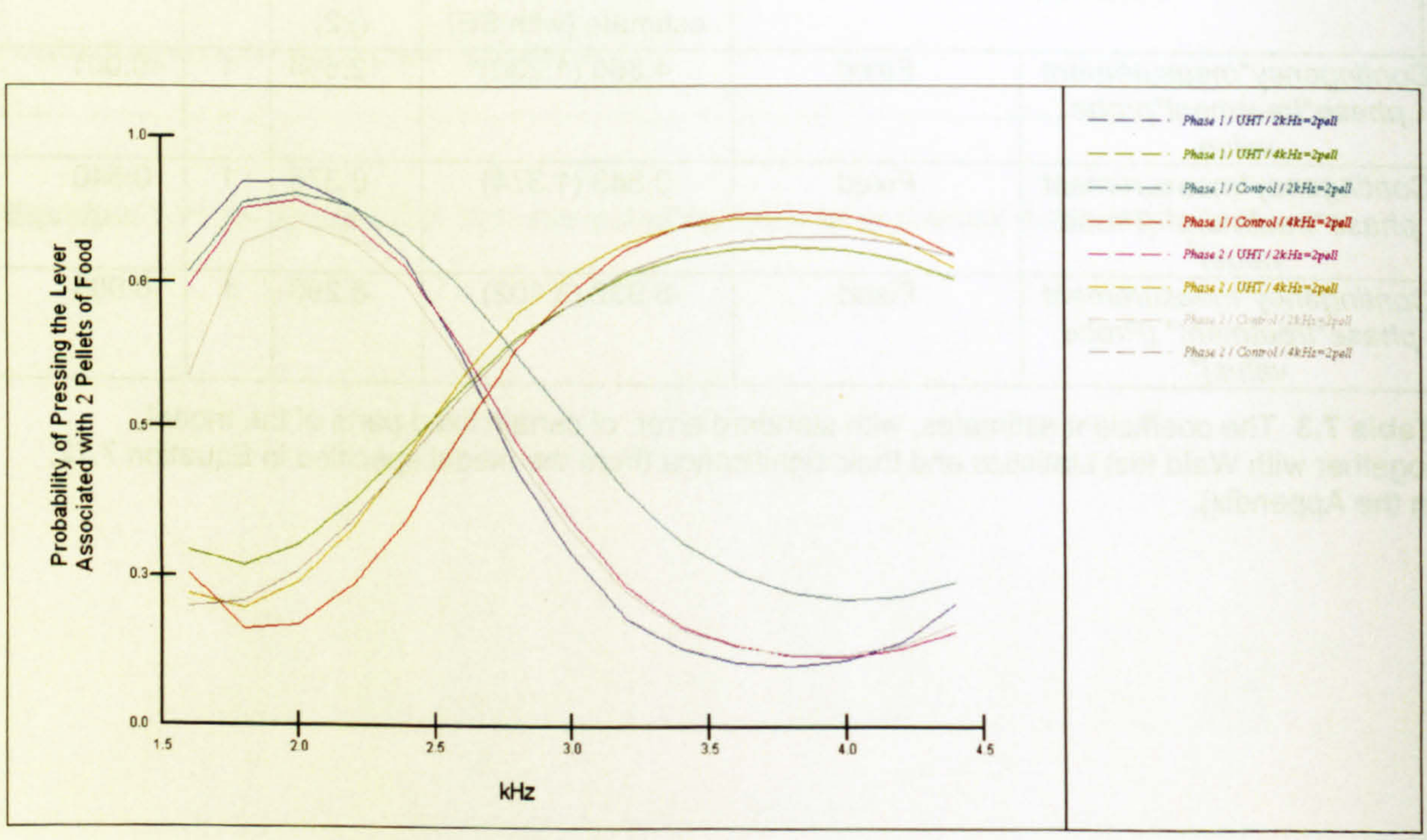

Figure 7.16 As Figure 7.15, but with kHz on the x-axis.

2-pell lvr on (1=press)$_{ij}$ ~ Binomial(Constant$_{ij}$, π_{ij})

logit(π_{ij}) = β_{0j}Constant + β_{1j}Standard log scale$_{ij}$ + β_{2j}Standard log scale quadratic$_{ij}$ + -8.099(0.570)Standard log scale cubic$_{ij}$ + 0.308(0.222)4kHz=2pell$_j$ +
-2.031(0.413)4kHz=2pell.Standard log scale$_{ij}$ + -2.685(0.551)4kHz=2pell.Standard log scale quadratic$_{ij}$ +
3.345(0.772)4kHz=2pell.Standard log scale cubic$_{ij}$ + -0.672(0.107)Phase2$_{ij}$ + 0.706(0.148)Phase2.4kHz=2pell$_{ij}$ + -0.518(0.212)UHT$_j$ +
0.528(0.287)UHT.4kHz=2pell$_j$ + 0.672(0.155)UHT.Phase2$_{ij}$ + -0.642(0.211)UHT.4kHz=2pell.Phase2$_{ij}$ + 0.009(0.005)Latency (secs)$_{ij}$

β_{0j} = 0.355(0.160) + u_{0j}

β_{1j} = 6.009(0.304) + u_{1j}

β_{2j} = 0.905(0.393) + u_{2j}

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) \quad : \quad \Omega_u = \begin{bmatrix} 0.080(0.036) & & \\ -0.110(0.056) & 0.307(0.133) & \\ -0.049(0.086) & 0.378(0.184) & 0.739(0.375) \end{bmatrix}$$

var(2-pell lvr on (1=press)$_{ij}$| π_{ij}) = π_{ij}(1 - π_{ij})/Constant$_{ij}$

Equation 7.13 The coefficient estimates generated by MLwiN for the specified model (modelling data from both *measurement phases*).

2-pell lvr on (1=press)$_{ij}$ ~ Binomial(Constant$_{ij}$, π_{ij})

logit(π_{ij}) = β_{0j}Constant + β_{1j}Standard log scale$_{ij}$ + β_{2j}Standard log scale quadratic$_{ij}$ + -8.407(0.618)Standard log scale cubic$_{ij}$ + 0.342(0.225)4kHz=2pell$_j$ +
-2.360(0.403)4kHz=2pell.Standard log scale$_{ij}$ + -2.718(0.548)4kHz=2pell.Standard log scale quadratic$_{ij}$ +
3.461(0.822)4kHz=2pell.Standard log scale cubic$_{ij}$ + -0.776(0.120)Phase2$_{ij}$ + 0.879(0.160)Phase2.4kHz=2pell$_{ij}$ + -0.626(0.220)UHT$_j$ +
0.587(0.292)UHT.4kHz=2pell$_j$ + 0.848(0.167)UHT.Phase2$_{ij}$ + -0.901(0.224)UHT.4kHz=2pell.Phase2$_{ij}$

β_{0j} = 0.356(0.166) + u_{0j}

β_{1j} = 6.506(0.301) + u_{1j}

β_{2j} = 0.756(0.397) + u_{2j}

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) \quad : \quad \Omega_u = \begin{bmatrix} 0.079(0.036) & & \\ -0.095(0.050) & 0.226(0.106) & \\ -0.050(0.084) & 0.397(0.170) & 0.658(0.354) \end{bmatrix}$$

var(2-pell lvr on (1=press)$_{ij}$| π_{ij}) = π_{ij}(1 - π_{ij})/Constant$_{ij}$

Equation 7.14 The coefficient estimates generated by MLwiN for the specified model, modelling only those lever presses occurring within three seconds.

Parameter		Coefficient estimate (with SE)	Wald ($\chi 2$)	Df	*P*
Intercept	Random at *subject* level	0.079 (0.036)	4.849	1	0.028 *
Probe value	Random at *subject* level	0.226 (0.106)	4.541	1	0.033 *
	Fixed	6.506 (0.301)	468.214	1	<0.001 **
(Probe value)^2	Fixed	0.756 (0.397)	3.631	1	0.057
(Probe value)^3	Fixed	-8.407 (0.618)	185.155	1	<0.001 **
Contingency	Fixed	0.342 (0.225)	2.297	1	0.130
*Contingency*Probe value*	Fixed	-2.360 (0.403)	34.281	1	<0.001 **
Contingency(Probe value)^2*	Fixed	-2.718 (0.548)	24.586	1	<0.001 **
Contingency(Probe value)^3*	Fixed	3.461 (0.822)	17.725	1	<0.001 **
Measurement phase	Fixed	-0.776 (0.120)	42.105	1	<0.001 **
*Measurement phase *Contingency*	Fixed	0.879 (0.160)	30.366	1	<0.001 **
Treatment	Fixed	-0.626 (0.220)	8.094	1	0.004 **
*Treatment*Contingency*	Fixed	0.587 (0.292)	4.030	1	0.047 *
*Treatment*Measurement phase*	Fixed	0.848 (0.167)	25.654	1	<0.001 **
*Treatment*Contingency *Measurement phase*	Fixed	-0.901 (0.224)	16.161	1	<0.001 **

Table 7.4 The coefficient estimates, with standard error, of fixed and random parts of the model (modelling only those lever responses made within 3 seconds), together with Wald test statistics and their significance (from the model specified in Equation 7.14, in the Appendix).

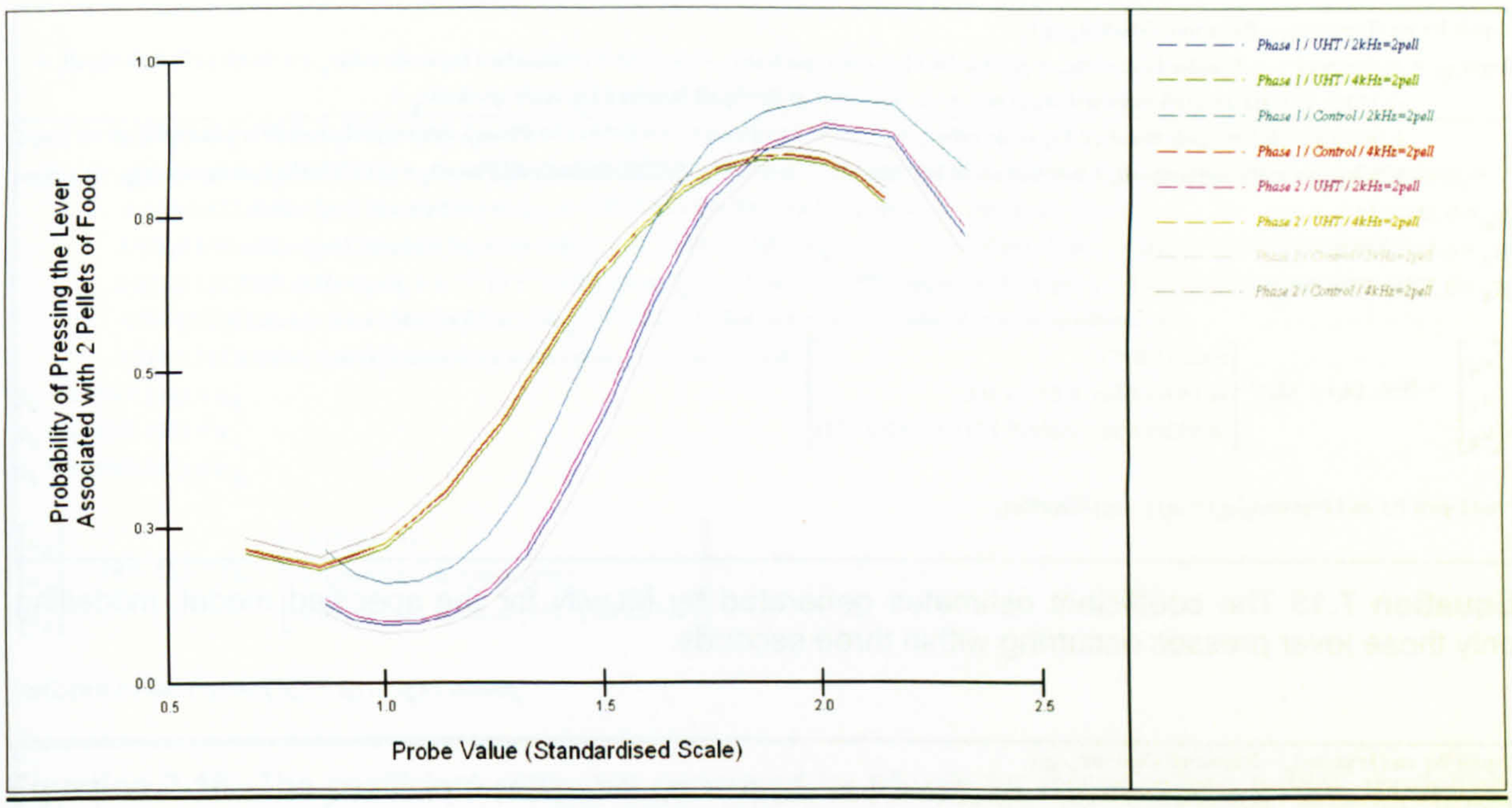

Figure 7.17 The probability of pressing the lever associated with 2 pellets of food (as opposed to the lever associated with 1 pellet of food) across different values of the probe stimuli (a standardised scale, with *1.0* corresponding to the reference tone associated with 1 pellet of food, and *2.0* corresponding to the reference tone associated with 2 pellets of food), by *measurement phase / treatment / contingency* group. These predicted lines were generated from the model and estimates in Equation 7.14; only those lever responses recorded withing 3 seconds are included.

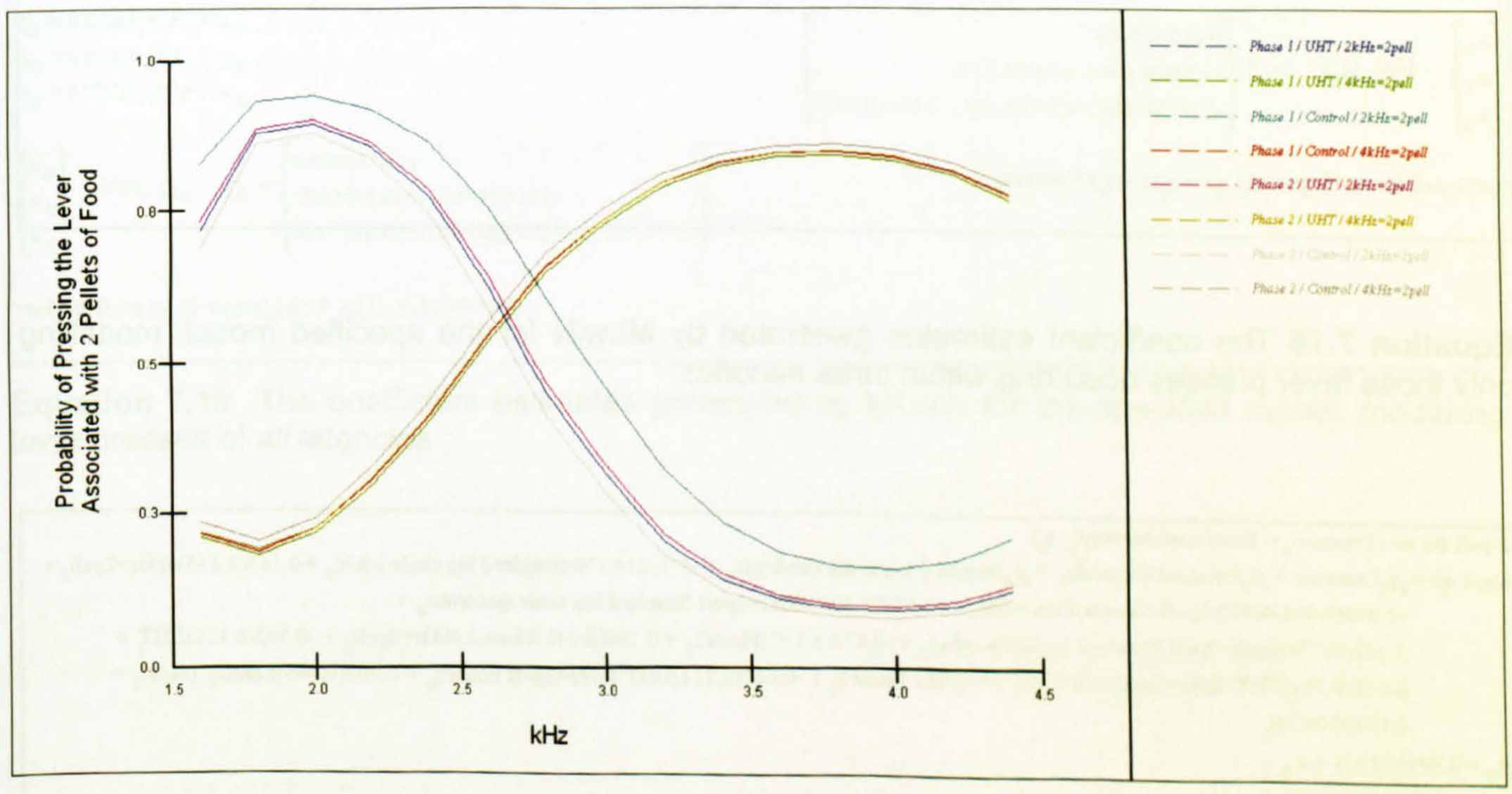

Figure 7.18 As Figure 7.17, but with kHz on the x-axis.

2-pell lvr on (1=press)$_{ij}$ ~ Binomial(Constant$_{ij}$, π_{ij})

logit(π_{ij}) = β_{0j}Constant + β_{1j}Standard log scale$_{ij}$ + β_{2j}Standard log scale quadratic$_{ij}$ + -8 512(0 619)Standard log scale cubic$_{ij}$ + 0 363(0 227)4kHz=2pell$_j$ + -2.354(0.403)4kHz=2pell.Standard log scale$_{ij}$ + -2 743(0 558)4kHz=2pell Standard log scale quadratic$_{ij}$ + 3 565(0.823)4kHz=2pell.Standard log scale cubic$_{ij}$ + -0.783(0 120)Phase2$_{ij}$ + 0 868(0 160)Phase2.4kHz=2pell$_{ij}$ + -0 591(0.223)UHT$_j$ + 0 555(0 295)UHT.4kHz=2pell$_j$ + 0.850(0 168)UHT Phase2$_{ij}$ + -0 875(0 225)UHT.4kHz=2pell Phase2$_{ij}$ + 0 197(0 053)Latency (secs)$_{ij}$

β_{0j} = 0.344(0.167) + u_{0j}

β_{1j} = 6.543(0.301) + u_{1j}

β_{2j} = 0.728(0.404) + u_{2j}

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0.081(0\ 037) & & \\ -0\ 094(0.050) & 0\ 225(0.106) & \\ -0\ 042(0\ 086) & 0.403(0\ 173) & 0\ 702(0.372) \end{bmatrix}$$

var(2-pell lvr on (1=press)$_{ij}$|π_{ij}) = π_{ij}(1 - π_{ij})/Constant$_{ij}$

Equation 7.15 The coefficient estimates generated by MLwiN for the specified model, modelling only those lever presses occurring within three seconds.

2-pell lvr on (1=press)$_{ij}$ ~ Binomial(Constant$_{ij}$, π_{ij})

logit(π_{ij}) = β_{0j}Constant + β_{1j}Standard log scale$_{ij}$ + β_{2j}Standard log scale quadratic$_{ij}$ + -8 257(0 649)Standard log scale cubic$_{ij}$ + 0 362(0 227)4kHz=2pell$_j$ + -2 425(0.430)4kHz=2pell.Standard log scale$_{ij}$ + -2 873(0 636)4kHz=2pell Standard log scale quadratic$_{ij}$ + 3 502(0 858)4kHz=2pell Standard log scale cubic$_{ij}$ + -0 726(0 122)Phase2$_{ij}$ + 0 805(0 163)Phase2.4kHz=2pell$_{ij}$ + -0 726(0.215)UHT$_j$ + 0.546(0.285)UHT 4kHz=2pell$_j$ + 0.772(0.171)UHT Phase2$_{ij}$ + -0.752(0 229)UHT 4kHz=2pell Phase2$_{ij}$ + 0 062(0 068)Latency (secs)$_{ij}$ + -2.915(0.234)Latency (secs).Standard log scale$_{ij}$ + 0.291(0 241)Latency (secs).Standard log scale quadratic$_{ij}$ + 3 245(0.514)Latency (secs) Standard log scale cubic$_{ij}$

β_{0j} = 0 394(0.167) + u_{0j}

β_{1j} = 6 689(0.321) + u_{1j}

β_{2j} = 0.580(0.462) + u_{2j}

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0\ 085(0\ 038) & & \\ -0.128(0\ 060) & 0.292(0.134) & \\ -0\ 008(0\ 100) & 0.169(0\ 187) & 1\ 033(0\ 502) \end{bmatrix}$$

var(2-pell lvr on (1=press)$_{ij}$|π_{ij}) = π_{ij}(1 - π_{ij})/Constant$_{ij}$

Equation 7.16 The coefficient estimates generated by MLwiN for the specified model, modelling only those lever presses occurring within three seconds.

2-pell lvr on (1=press)$_{ij}$ ~ Binomial(Constant$_{ij}$, π_{ij})

logit(π_{ij}) = β_{0j}Constant + β_{1j}Standard log scale$_{ij}$ + β_{2j}Standard log scale quadratic$_{ij}$ + -8 124(0 570)Standard log scale cubic$_{ij}$ + 0 265(0.135)4kHz=2pell$_j$ + -2 034(0.411)4kHz=2pell.Standard log scale$_{ij}$ + -2.672(0 549)4kHz=2pell Standard log scale quadratic$_{ij}$ + 3 341(0.770)4kHz=2pell.Standard log scale cubic$_{ij}$ + -0 672(0.107)Phase2$_{ij}$ + 0.706(0.148)Phase2 4kHz=2pell$_{ij}$ + -0.542(0 120)UHT$_j$ + 0 618(0 164)UHT.4kHz=2pell$_j$ + 0 672(0.155)UHT Phase2$_{ij}$ + -0 642(0 211)UHT.4kHz=2pell Phase2$_{ij}$ + 0 008(0.005)Latency (secs)$_{ij}$ + 0 534(0.063)R$_j$

β_{0j} = 0 099(0.103) + u_{0j}

β_{1j} = 6 009(0.302) + u_{1j}

β_{2j} = 0 921(0.391) + u_{2j}

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0\ 012(0.012) & & \\ -0\ 059(0\ 031) & 0.303(0\ 131) & \\ -0\ 079(0.058) & 0\ 387(0.184) & 0.729(0\ 370) \end{bmatrix}$$

var(2-pell lvr on (1=press)$_{ij}$|π_{ij}) = π_{ij}(1 - π_{ij})/Constant$_{ij}$

Equation 7.17 The coefficient estimates generated by MLwiN for the specified model, modelling lever presses of *all* latencies.

2-pell lvr on (1=press)$_{ij}$ ~ Binomial(Constant$_{ij}$, π_{ij})

logit(π_{ij}) = β_{0j}Constant + β_{1j}Standard log scale$_{ij}$ + β_{2j}Standard log scale quadratic$_{ij}$ + -8.298(0 651)Standard log scale cubic$_{ij}$ + 0.319(0 151)4kHz=2pell$_j$ + -2.453(0 425)4kHz=2pell.Standard log scale$_{ij}$ + -2 856(0 644)4kHz=2pell.Standard log scale quadratic$_{ij}$ + 3 558(0 858)4kHz=2pell.Standard log scale cubic$_{ij}$ + -0 732(0 122)Phase2$_{ij}$ + 0.812(0.163)Phase2.4kHz=2pell$_{ij}$ + -0 594(0 145)UHT$_j$ + 0 623(0.192)UHT.4kHz=2pell$_j$ + 0 [illegible](0 171)UHT.Phase2$_{ij}$ + -0 [illegible]59(0 229)UHT.4kHz=2pell.Phase2$_{ij}$ + 0 055(0 06[illegible])Latency (secs)$_{ij}$ + -2 893(0 234)Latency (secs) Standard log scale$_{ij}$ + 0 304(0 240)Latency (secs) Standard log scale quadratic$_{ij}$ + 3 214(0.513)Latency (secs) Standard log scale cubic$_{ij}$ + 0.454(0.0[illegible]6)R$_j$

β_{0j} = 0 106(0 118) + u_{0j}

β_{1j} = 6.702(0.318) + u_{1j}

β_{2j} = 0 600(0 467) + u_{2j}

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0\ 015(0\ 014) & & \\ -0.059(0\ 033) & 0\ 278(0\ 129) & \\ -0.05[illegible](0\ 066) & 0\ 180(0\ 187) & 1\ 075(0\ 517) \end{bmatrix}$$

var(2-pell lvr on (1=press)$_{ij}$|π_{ij}) = $\pi_{ij}(1 - \pi_{ij})$/Constant$_{ij}$

Equation 7.18 The coefficient estimates generated by MLwiN for the specified model, modelling only those lever presses occurring within three seconds.

2-pell lvr on (1=press)$_{ij}$ ~ Binomial(Constant$_{ij}$, π_{ij})

logit(π_{ij}) = β_{0j}Constant + β_{1j}Standard log scale$_{ij}$ + β_{2j}Standard log scale quadratic$_{ij}$ + -8 125(0 570)Standard log scale cubic$_{ij}$ + 0 261(0.131)4kHz=2pell$_j$ + -2.041(0 411)4kHz=2pell.Standard log scale$_{ij}$ + -2 670(0 551)4kHz=2pell.Standard log scale quadratic$_{ij}$ + 3.358(0 770)4kHz=2pell Standard log scale cubic$_{ij}$ + -0 672(0.107)Phase2$_{ij}$ + 0 706(0.148)Phase2.4kHz=2pell$_{ij}$ + -0 548(0 118)UHT$_j$ + 0 625(0 161)UHT.4kHz=2pell$_j$ + 0 672(0 155)UHT Phase2$_{ij}$ + -0 642(0 211)UHT.4kHz=2pell Phase2$_{ij}$ + 0 008(0 005)Latency (secs)$_{ij}$ + 0 597(0 086)R$_j$ + 0.099(0.084)B$_j$ + -0 137(0 121)R.B$_j$

β_{0j} = 0.056(0 108) + u_{0j}

β_{1j} = 6 013(0.302) + u_{1j}

β_{2j} = 0 922(0 392) + u_{2j}

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0\ 009(0\ 011) & & \\ -0\ 058(0\ 030) & 0.305(0.131) & \\ -0.070(0\ 055) & 0.390(0\ 185) & 0.737(0.374) \end{bmatrix}$$

var(2-pell lvr on (1=press)$_{ij}$|π_{ij}) = $\pi_{ij}(1 - \pi_{ij})$/Constant$_{ij}$

Equation 7.19 The coefficient estimates generated by MLwiN for the specified model, modelling lever presses of *all* latencies.

2-pell lvr on (1=press)$_{ij}$ ~ Binomial(Constant$_{ij}$, π_{ij})

logit(π_{ij}) = β_{0j}Constant + β_{1j}Standard log scale$_{ij}$ + β_{2j}Standard log scale quadratic$_{ij}$ + -8 299(0 651)Standard log scale cubic$_{ij}$ + 0 300(0 144)4kHz=2pell$_j$ +
-2.450(0 427)4kHz=2pell.Standard log scale$_{ij}$ + -2 843(0 643)4kHz=2pell Standard log scale quadratic$_{ij}$ +
3 549(0 858)4kHz=2pell Standard log scale cubic$_{ij}$ + -0 732(0 122)Phase2$_{ij}$ + 0 811(0 163)Phase2 4kHz=2pell$_{ij}$ + -0 566(0 140)UHT$_j$ +
0 663(0 188)UHT 4kHz=2pell$_j$ + 0 777(0 171)UHT Phase2$_{ij}$ + -0 759(0 229)UHT.4kHz=2pell Phase2$_{ij}$ + 0 060(0 068)Latency (secs)$_{ij}$ +
-2.893(0 234)Latency (secs).Standard log scale$_{ij}$ + 0 286(0 241)Latency (secs).Standard log scale quadratic$_{ij}$ +
3 221(0.513)Latency (secs).Standard log scale cubic$_{ij}$ + 0 537(0 106)R$_j$ + 0 115(0 104)B$_j$ + -0 122(0 149)R.B$_j$

β_{0j} = 0 023(0.125) + u_{0j}
β_{1j} = 6 702(0 319) + u_{1j}
β_{2j} = 0 603(0 466) + u_{2j}

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0\ 009(0\ 011) & & \\ -0\ 036(0.029) & 0\ 284(0\ 131) & \\ -0\ 053(0\ 062) & 0\ 183(0\ 188) & 1\ 064(0\ 514) \end{bmatrix}$$

var(2-pell lvr on (1=press)$_{ij}$|π_{ij}) = π_{ij}(1 - π_{ij})/Constant$_{ij}$

Equation 7.20 The coefficient estimates generated by MLwiN for the specified model, modelling lever presses of *all* latencies.

$$\text{Latency}_{ij} = -1.364(0.005) + e_{ij}$$

$$e_{ij} \sim N(0, \sigma_e^2) \quad \sigma_e^2 = 0.247(0.004)$$

$$-2*loglikelihood = 12423.940(8619 \text{ of } 8619 \text{ cases in use})$$

Equation 7.21 The coefficient estimates generated by MLwiN for the specified model, with latency as the response (*y*) variable.

$$\text{Latency}_{ij} = \beta_{0j} + e_{ij}$$

$$\beta_{0j} = -1.364(0.052) + u_{0j}$$

$$u_{0j} \sim N(0, \sigma_{u0}^2) \quad \sigma_{u0}^2 = 0.043(0.015)$$

$$e_{ij} \sim N(0, \sigma_e^2) \quad \sigma_e^2 = 0.204(0.003)$$

$$-2*loglikelihood = 10839.530(8619 \text{ of } 8619 \text{ cases in use})$$

Equation 7.22 The coefficient estimates generated by MLwiN for the specified model, with latency as the response (*y*) variable.

Latency$_{ij}$ = β_{0j} + -0.150(0 025)Standard log scale$_{ij}$ + 0.289(0 026)Standard log scale quadratic$_{ij}$ + 0.225(0 059)Standard log scale cubic$_{ij}$ + e_{ij}
β_{0j} = -1.421(0 053) + u_{0j}

$u_{0j} \sim N(0, \sigma_{u0}^2)$ σ_{u0}^2 = 0 044(0.015)
$e_{ij} \sim N(0, \sigma_e^2)$ σ_e^2 = 0.200(0.003)
-2*loglikelihood = 10665.090(8619 of 8619 cases in use)

Equation 7.23 The coefficient estimates generated by MLwiN for the specified model, with latency as the response (*y*) variable.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + 0.267(0.064)\text{Standard log scale cubic}_{ij} + e_{ij}$$

$$\beta_{0j} = -1.410(0.052) + u_{0j}$$

$$\beta_{1j} = -0.161(0.042) + u_{1j}$$

$$\beta_{2j} = 0.242(0.042) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0.043(0.015) & & \\ -0.006(0.007) & 0.018(0.007) & \\ 0.000(0.009) & -0.002(0.006) & 0.017(0.010) \end{bmatrix}$$

$$e_{ij} \sim N(0, \sigma_e^2) \quad \sigma_e^2 = 0.196(0.003)$$

$$-2 * loglikelihood = 10544.000(8619 \text{ of } 8619 \text{ cases in use})$$

Equation 7.24 The coefficient estimates generated by MLwiN for the specified model, with latency as the response (y) variable.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + 0.360(0\ 066)\text{Standard log scale cubic}_{ij} + 0\ 076(0.011)\text{2pell lvr press}_{ij} + e_{ij}$$

$$\beta_{0j} = -1\ 453(0\ 053) + u_{0j}$$

$$\beta_{1j} = -0.236(0\ 043) + u_{1j}$$

$$\beta_{2j} = 0\ 255(0\ 045) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0\ 043(0\ 015) & & \\ -0\ 005(0\ 007) & 0\ 017(0\ 007) & \\ 0\ 001(0.009) & -0\ 002(0.006) & 0\ 021(0\ 011) \end{bmatrix}$$

$$e_{ij} \sim N(0, \sigma_e^2) \quad \sigma_e^2 = 0\ 195(0\ 003)$$

$$-2 * loglikelihood = 10500.380(8619 \text{ of } 8619 \text{ cases in use})$$

Equation 7.25 The coefficient estimates generated by MLwiN for the specified model, with latency as the response (y) variable.

$$\begin{aligned}\text{Latency}_{ij} = {} & \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -0.337(0\ 092)\text{Standard log scale cubic}_{ij} + 0\ 037(0\ 015)\text{2pell lvr press}_{ij} + \\ & -0.944(0\ 052)\text{2pell lvr press Standard log scale}_{ij} + 0.130(0\ 056)\text{2pell lvr press.Standard log scale quadratic}_{ij} + \\ & 0.998(0\ 120)\text{2pell lvr press.Standard log scale cubic}_{ij} + e_{ij}\end{aligned}$$

$$\beta_{0j} = -1.388(0\ 052) + u_{0j}$$

$$\beta_{1j} = 0.317(0.048) + u_{1j}$$

$$\beta_{2j} = 0.313(0.050) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0.042(0\ 015) & & \\ -0\ 005(0\ 006) & 0\ 012(0\ 005) & \\ -0\ 002(0\ 008) & 0\ 003(0.005) & 0.014(0.009) \end{bmatrix}$$

$$e_{ij} \sim N(0, \sigma_e^2) \quad \sigma_e^2 = 0\ 183(0\ 003)$$

$$-2 * loglikelihood = 9960.326(8619 \text{ of } 8619 \text{ cases in use})$$

Equation 7.26 The coefficient estimates generated by MLwiN for the specified model, with latency as the response (y) variable.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -0\ 337(0\ 092)\text{Standard log scale cubic}_{ij} + 0\ 037(0.015)\text{2pell lvr press}_{ij} + -0\ 944(0.052)\text{2pell lvr press Standard log scale}_{ij} + 0\ 130(0\ 056)\text{2pell lvr press Standard log scale quadratic}_{ij} + 0\ 998(0\ 120)\text{2pell lvr press Standard log scale cubic}_{ij} + -0\ 004(0\ 100)\text{4kHz=2pell}_{j} + e_{ij}$$

$$\beta_{0j} = -1.386(0\ 072) + u_{0j}$$

$$\beta_{1j} = 0\ 317(0\ 048) + u_{1j}$$

$$\beta_{2j} = 0.313(0\ 050) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0,\ \Omega_u) : \Omega_u = \begin{bmatrix} 0\ 042(0.015) & & \\ -0\ 005(0.006) & 0\ 012(0\ 005) & \\ -0\ 002(0\ 008) & 0\ 003(0\ 005) & 0\ 014(0\ 009) \end{bmatrix}$$

$$e_{ij} \sim N(0,\ \sigma_e^2) \quad \sigma_e^2 = 0\ 183(0\ 003)$$

$$-2*\text{loglikelihood} = 9960.324(8619 \text{ of } 8619 \text{ cases in use})$$

Equation 7.27 The coefficient estimates generated by MLwiN for the specified model, with latency as the response (*y*) variable.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -0\ 258(0\ 153)\text{Standard log scale cubic}_{ij} + 0\ 105(0\ 022)\text{2pell lvr press}_{ij} + -1.249(0\ 085)\text{2pell lvr press Standard log scale}_{ij} + 0\ 136(0\ 098)\text{2pell lvr press.Standard log scale quadratic}_{ij} + 1\ 137(0.213)\text{2pell lvr press Standard log scale cubic}_{ij} + 0\ 084(0.103)\text{4kHz=2pell}_{j} + -0\ 206(0\ 101)\text{4kHz=2pell Standard log scale}_{ij} + -0\ 220(0.111)\text{4kHz=2pell Standard log scale quadratic}_{ij} + -0\ 207(0\ 213)\text{4kHz=2pell.Standard log scale cubic}_{ij} + -0\ 176(0.032)\text{4kHz=2pell 2pell lvr press}_{ij} + 0\ 646(0\ 115)\text{4kHz=2pell 2pell lvr press Standard log scale}_{ij} + 0.234(0\ 139)\text{4kHz=2pell.2pell lvr press.Standard log scale quadratic}_{ij} + -0\ 381(0\ 294)\text{4kHz=2pell 2pell lvr press.Standard log scale cubic}_{ij} + e_{ij}$$

$$\beta_{0j} = -1\ 407(0\ 072) + u_{0j}$$

$$\beta_{1j} = 0\ 404(0\ 075) + u_{1j}$$

$$\beta_{2j} = 0.313(0\ 066) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0,\ \Omega_u) : \Omega_u = \begin{bmatrix} 0\ 040(0\ 014) & & \\ -0\ 005(0\ 006) & 0\ 011(0\ 004) & \\ -0\ 003(0\ 007) & 0\ 004(0\ 004) & 0\ 011(0\ 007) \end{bmatrix}$$

$$e_{ij} \sim N(0,\ \sigma_e^2) \quad \sigma_e^2 = 0\ 180(0\ 003)$$

$$-2*\text{loglikelihood} = 9814\ 120(8619 \text{ of } 8619 \text{ cases in use})$$

Equation 7.28 The coefficient estimates generated by MLwiN for the specified model, with latency as the response (*y*) variable.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -0\ 257(0.153)\text{Standard log scale cubic}_{ij} + 0.106(0\ 022)\text{2pell lvr press}_{ij} + -1\ 249(0.085)\text{2pell lvr press.Standard log scale}_{ij} + 0.137(0\ 098)\text{2pell lvr press Standard log scale quadratic}_{ij} + 1\ 138(0.213)\text{2pell lvr press.Standard log scale cubic}_{ij} + 0\ 085(0\ 103)\text{4kHz=2pell}_{j} + -0\ 205(0\ 101)\text{4kHz=2pell Standard log scale}_{ij} + -0.220(0\ 111)\text{4kHz=2pell Standard log scale quadratic}_{ij} + -0.208(0.213)\text{4kHz=2pell.Standard log scale cubic}_{ij} + -0.177(0\ 032)\text{4kHz=2pell.2pell lvr press}_{ij} + 0\ 647(0.115)\text{4kHz=2pell.2pell lvr press.Standard log scale}_{ij} + 0.234(0.139)\text{4kHz=2pell 2pell lvr press Standard log scale quadratic}_{ij} + -0\ 385(0\ 294)\text{4kHz=2pell 2pell lvr press.Standard log scale cubic}_{ij} + 0\ 013(0\ 009)\text{Phase2}_{ij} + e_{ij}$$

$$\beta_{0j} = -1\ 414(0.072) + u_{0j}$$

$$\beta_{1j} = 0\ 403(0\ 075) + u_{1j}$$

$$\beta_{2j} = 0.312(0\ 066) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0,\ \Omega_u) : \Omega_u = \begin{bmatrix} 0\ 040(0\ 014) & & \\ -0\ 005(0\ 006) & 0\ 011(0\ 004) & \\ -0.003(0\ 007) & 0\ 004(0\ 004) & 0\ 011(0\ 007) \end{bmatrix}$$

$$e_{ij} \sim N(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.180(0.003)$$

$$-2*\text{loglikelihood} = 9811.991(8619 \text{ of } 8619 \text{ cases in use})$$

Equation 7.29 The coefficient estimates generated by MLwiN for the specified model, with latency as the response (*y*) variable.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + \text{-0 254(0 153)Standard log scale cubic}_{ij} + \text{0 135(0 024)2pell lvr press}_{ij} +$$
$$\text{-1 244(0.084)2pell lvr press Standard log scale}_{ij} + \text{0 137(0 098)2pell lvr press.Standard log scale quadratic}_{ij} +$$
$$\text{1 132(0 213)2pell lvr press.Standard log scale cubic}_{ij} + \text{0.087(0 103)4kHz=2pell}_{j} + \text{-0.202(0.101)4kHz=2pell Standard log scale}_{ij} +$$
$$\text{-0 219(0 111)4kHz=2pell Standard log scale quadratic}_{ij} + \text{-0 210(0 213)4kHz=2pell Standard log scale cubic}_{ij} +$$
$$\text{-0 177(0 032)4kHz=2pell.2pell lvr press}_{ij} + \text{0 641(0.115)4kHz=2pell 2pell lvr press Standard log scale}_{ij} +$$
$$\text{0.233(0.139)4kHz=2pell.2pell lvr press.Standard log scale quadratic}_{ij} + \text{-0 376(0.293)4kHz=2pell.2pell lvr press Standard log scale cubic}_{ij} +$$
$$\text{0 044(0.013)Phase2}_{ij} + \text{-0.058(0 018)Phase2.2pell lvr press}_{ij} + e_{ij}$$

$$\beta_{0j} = \text{-1.431(0.073)} + u_{0j}$$
$$\beta_{1j} = \text{0.400(0 075)} + u_{1j}$$
$$\beta_{2j} = \text{0.311(0.066)} + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} \text{0 040(0 014)} & & \\ \text{-0 005(0 006)} & \text{0 011(0.004)} & \\ \text{-0.003(0.007)} & \text{0 004(0 004)} & \text{0 011(0 007)} \end{bmatrix}$$

$$e_{ij} \sim N(0,\ \sigma_e^2) \quad \sigma_e^2 = \text{0 180(0 003)}$$

*-2*loglikelihood* = 9802 133(8619 of 8619 cases in use)

Equation 7.30 The coefficient estimates generated by MLwiN for the specified model, with latency as the response (*y*) variable.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + \text{-0.251(0 153)Standard log scale cubic}_{ij} + \text{0 132(0 024)2pell lvr press}_{ij} +$$
$$\text{-1 245(0.084)2pell lvr press Standard log scale}_{ij} + \text{0.140(0.098)2pell lvr press Standard log scale quadratic}_{ij} +$$
$$\text{1 134(0 212)2pell lvr press.Standard log scale cubic}_{ij} + \text{0 121(0 103)4kHz=2pell}_{j} + \text{-0 199(0 100)4kHz=2pell Standard log scale}_{ij} +$$
$$\text{-0.216(0.111)4kHz=2pell Standard log scale quadratic}_{ij} + \text{-0 214(0 213)4kHz=2pell Standard log scale cubic}_{ij} +$$
$$\text{-0 179(0.032)4kHz=2pell.2pell lvr press}_{ij} + \text{0 640(0 115)4kHz=2pell.2pell lvr press Standard log scale}_{ij} +$$
$$\text{0 228(0 138)4kHz=2pell 2pell lvr press Standard log scale quadratic}_{ij} + \text{-0 372(0 293)4kHz=2pell 2pell lvr press.Standard log scale cubic}_{ij} +$$
$$\text{0 073(0 016)Phase2}_{ij} + \text{-0 048(0 019)Phase2.2pell lvr press}_{ij} + \text{-0 067(0 019)Phase2 4kHz=2pell}_{ij} + e_{ij}$$

$$\beta_{0j} = \text{-1.446(0 073)} + u_{0j}$$
$$\beta_{1j} = \text{0 397(0 075)} + u_{1j}$$
$$\beta_{2j} = \text{0 309(0 066)} + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} \text{0 040(0 014)} & & \\ \text{-0 005(0 006)} & \text{0 011(0 004)} & \\ \text{-0 003(0 007)} & \text{0 004(0 004)} & \text{0 011(0 007)} \end{bmatrix}$$

$$e_{ij} \sim N(0,\ \sigma_e^2) \quad \sigma_e^2 = \text{0 180(0.003)}$$

*-2*loglikelihood* = 9788 931(8619 of 8619 cases in use)

Equation 7.31 The coefficient estimates generated by MLwiN for the specified model, with latency as the response (*y*) variable.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + \text{-0 119(0 165)Standard log scale cubic}_{ij} + \text{0 138(0 025)2pell lvt press}_{ij} +$$
$$\text{-1.237(0 085)2pell lvt press.Standard log scale}_{ij} + \text{0 135(0 098)2pell lvt press Standard log scale quadratic}_{ij} +$$
$$\text{1 113(0.213)2pell lvt press.Standard log scale cubic}_{ij} + \text{0 128(0 103)4kHz=2pell}_{ij} + \text{-0.194(0 101)4kHz=2pell.Standard log scale}_{ij} +$$
$$\text{-0 219(0 111)4kHz=2pell Standard log scale quadratic}_{ij} + \text{-0 227(0 213)4kHz=2pell Standard log scale cubic}_{ij} +$$
$$\text{-0 179(0 032)4kHz=2pell.2pell lvt press}_{ij} + \text{0 632(0.115)4kHz=2pell.2pell lvt press.Standard log scale}_{ij} +$$
$$\text{0.231(0 138)4kHz=2pell.2pell lvt press.Standard log scale quadratic}_{ij} + \text{-0 350(0 293)4kHz=2pell 2pell lvt press Standard log scale cubic}_{ij} +$$
$$\text{0 095(0 020)Phase2}_{ij} + \text{-0 057(0 022)Phase2.2pell lvt press}_{ij} + \text{-0 079(0 019)Phase2.4kHz=2pell}_{ij} + \text{0.092(0 051)Phase2.Standard log scale}_{ij} +$$
$$\text{-0 058(0 049)Phase2.Standard log scale quadratic}_{ij} + \text{-0 236(0 114)Phase2 Standard log scale cubic}_{ij} + e_{ij}$$

$$\beta_{0j} = \text{-1.458(0 073)} + u_{0j}$$
$$\beta_{1j} = \text{0 347(0 080)} + u_{1j}$$
$$\beta_{2j} = \text{0.341(0.071)} + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} \text{0 040(0 014)} & & \\ \text{-0 005(0 006)} & \text{0 011(0 004)} & \\ \text{-0 003(0.007)} & \text{0 004(0 004)} & \text{0 011(0 007)} \end{bmatrix}$$

$$e_{ij} \sim N(0, \sigma_e^2) \quad \sigma_e^2 = \text{0 180(0 003)}$$

*-2*loglikelihood* = 9783.341(8619 of 8619 cases in use)

Equation 7.32 The coefficient estimates generated by MLwiN for the specified model, with latency as the response (*y*) variable.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + \text{-0 244(0.153)Standard log scale cubic}_{ij} + \text{0 132(0.024)2pell lvt press}_{ij} +$$
$$\text{-1 242(0.084)2pell lvt press.Standard log scale}_{ij} + \text{0 140(0 098)2pell lvt press Standard log scale quadratic}_{ij} +$$
$$\text{1 120(0.212)2pell lvt press Standard log scale cubic}_{ij} + \text{0 121(0 088)4kHz=2pell}_{ij} + \text{-0 198(0.101)4kHz=2pell Standard log scale}_{ij} +$$
$$\text{-0 218(0 111)4kHz=2pell.Standard log scale quadratic}_{ij} + \text{-0 221(0 213)4kHz=2pell Standard log scale cubic}_{ij} +$$
$$\text{-0 178(0 032)4kHz=2pell.2pell lvt press}_{ij} + \text{0 637(0.115)4kHz=2pell 2pell lvt press Standard log scale}_{ij} +$$
$$\text{0.227(0.138)4kHz=2pell.2pell lvt press Standard log scale quadratic}_{ij} + \text{-0.355(0 293)4kHz=2pell 2pell lvt press Standard log scale cubic}_{ij} +$$
$$\text{0 073(0.016)Phase2}_{ij} + \text{-0.048(0 019)Phase2.2pell lvt press}_{ij} + \text{-0 067(0 019)Phase2.4kHz=2pell}_{ij} + \text{-0 263(0 077)UHT}_{j} + e_{ij}$$

$$\beta_{0j} = \text{-1 315(0 073)} + u_{0j}$$
$$\beta_{1j} = \text{0.396(0 075)} + u_{1j}$$
$$\beta_{2j} = \text{0 311(0.066)} + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} \text{0 029(0.010)} & & \\ \text{-0 003(0 005)} & \text{0 011(0.004)} & \\ \text{-0.009(0 007)} & \text{0.004(0 004)} & \text{0 010(0.007)} \end{bmatrix}$$

$$e_{ij} \sim N(0, \sigma_e^2) \quad \sigma_e^2 = \text{0 180(0 003)}$$

*-2*loglikelihood* = 9781.123(8619 of 8619 cases in use)

Equation 7.33 The coefficient estimates generated by MLwiN for the specified model, with latency as the response (*y*) variable.

$$
\begin{aligned}
\text{Latency}_{ij} = {} & \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + \text{-0 245(0.153)Standard log scale cubic}_{ij} + \text{0 135(0 024)2pell lvr press}_{ij} + \\
& \text{-1.246(0.084)2pell lvr press Standard log scale}_{ij} + \text{0 134(0 098)2pell lvr press.Standard log scale quadratic}_{ij} + \\
& \text{1 129(0 212)2pell lvr press.Standard log scale cubic}_{ij} + \text{0 121(0 088)4kHz=2pell}_{j} + \text{-0 202(0.101)4kHz=2pell Standard log scale}_{ij} + \\
& \text{-0 218(0 111)4kHz=2pell Standard log scale quadratic}_{ij} + \text{-0 214(0 213)4kHz=2pell Standard log scale cubic}_{ij} + \\
& \text{-0 181(0.032)4kHz=2pell.2pell lvr press}_{ij} + \text{0 648(0.115)4kHz=2pell.2pell lvr press Standard log scale}_{ij} + \\
& \text{0 229(0 138)4kHz=2pell 2pell lvr press.Standard log scale quadratic}_{ij} + \text{-0.378(0.293)4kHz=2pell.2pell lvr press Standard log scale cubic}_{ij} + \\
& \text{0 100(0 018)Phase2}_{ij} + \text{-0 048(0.019)Phase2 2pell lvr press}_{ij} + \text{-0 067(0 018)Phase2.4kHz=2pell}_{ij} + \text{-0 236(0 077)UHT}_{j} + \\
& \text{-0 055(0 018)UHT Phase2}_{ij} + e_{ij} \\
\beta_{0j} = {} & \text{-1.330(0.073)} + u_{0j} \\
\beta_{1j} = {} & \text{0 396(0.075)} + u_{1j} \\
\beta_{2j} = {} & \text{0.313(0 066)} + u_{2j}
\end{aligned}
$$

$$
\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} \text{0 029(0 010)} & & \\ \text{-0 003(0 005)} & \text{0 011(0 004)} & \\ \text{-0 009(0.007)} & \text{0 004(0 004)} & \text{0 010(0.007)} \end{bmatrix}
$$

$e_{ij} \sim N(0,\ \sigma_e^2)\quad \sigma_e^2 = 0.180(0\ 003)$

*-2*loglikelihood* = 9772.200(8619 of 8619 cases in use)

Equation 7.34 The coefficient estimates generated by MLwiN for the specified model, with latency as the response (*y*) variable.

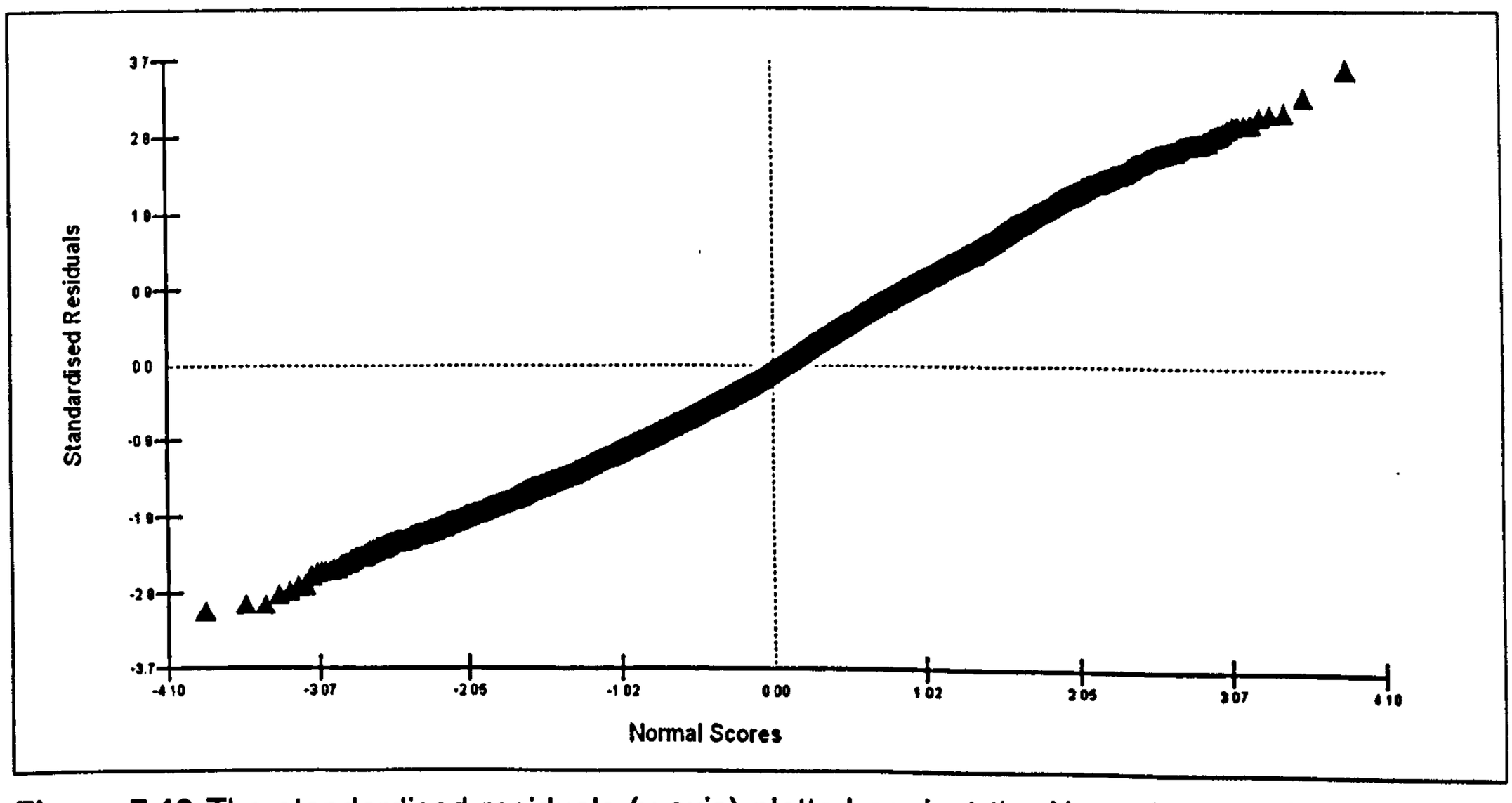

Figure 7.19 The standardised residuals (y-axis) plotted against the Normal scores for the Level 1 (i.e. *trial*) variance, from the model specified in Equation 7.34.

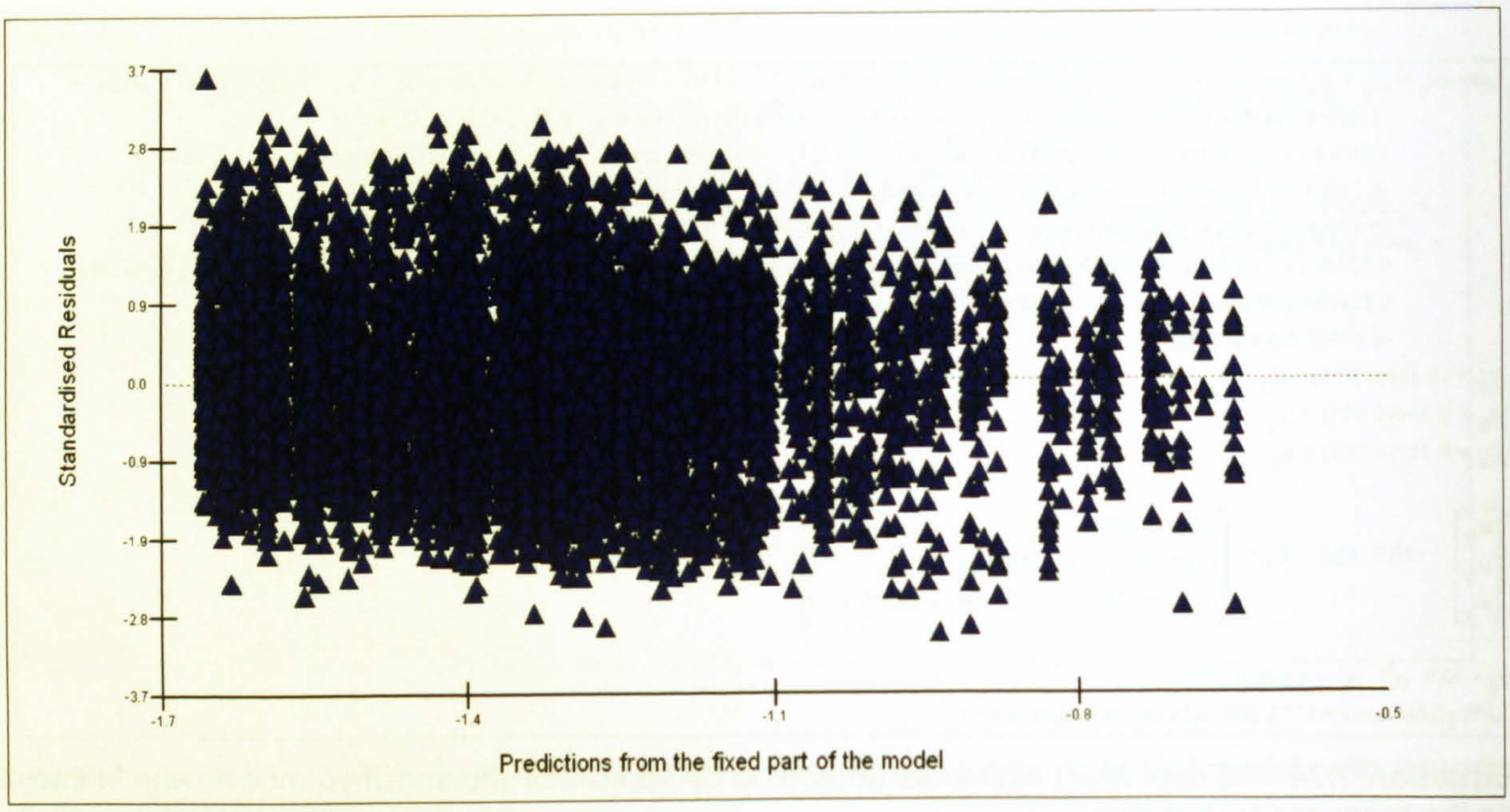

Figure 7.20 The standardised residuals (y-axis) plotted against the predicted values from the fixed part of the model (prior to backtransformation) for the Level 1 (i.e. *trial*) variance, from the model specified in Equation 7.34.

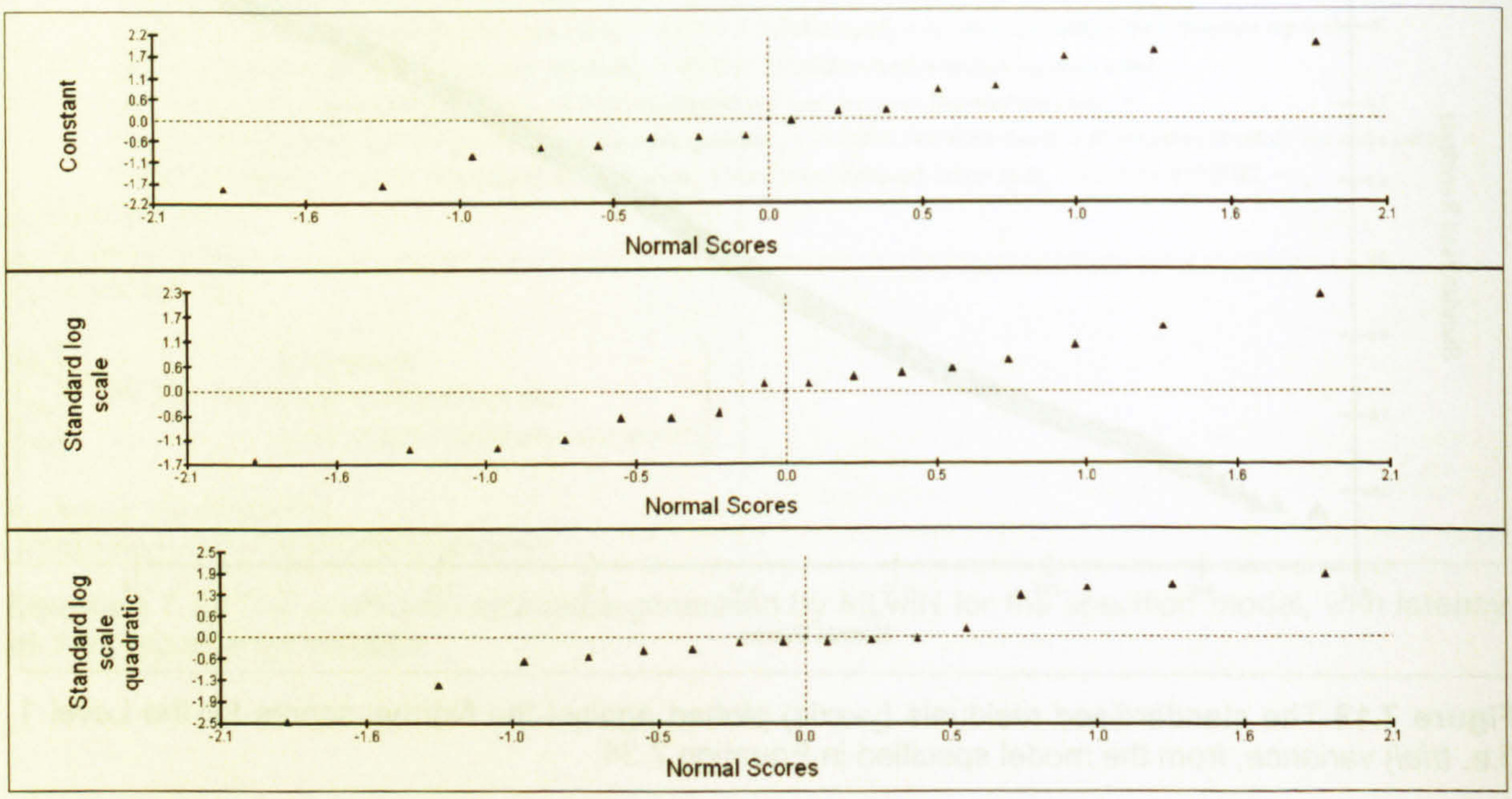

Figure 7.21 The standardised residuals (y-axis) plotted against the Normal scores for the Level 2 (i.e. *subject*) variance, from the model specified in Equation 7.34.

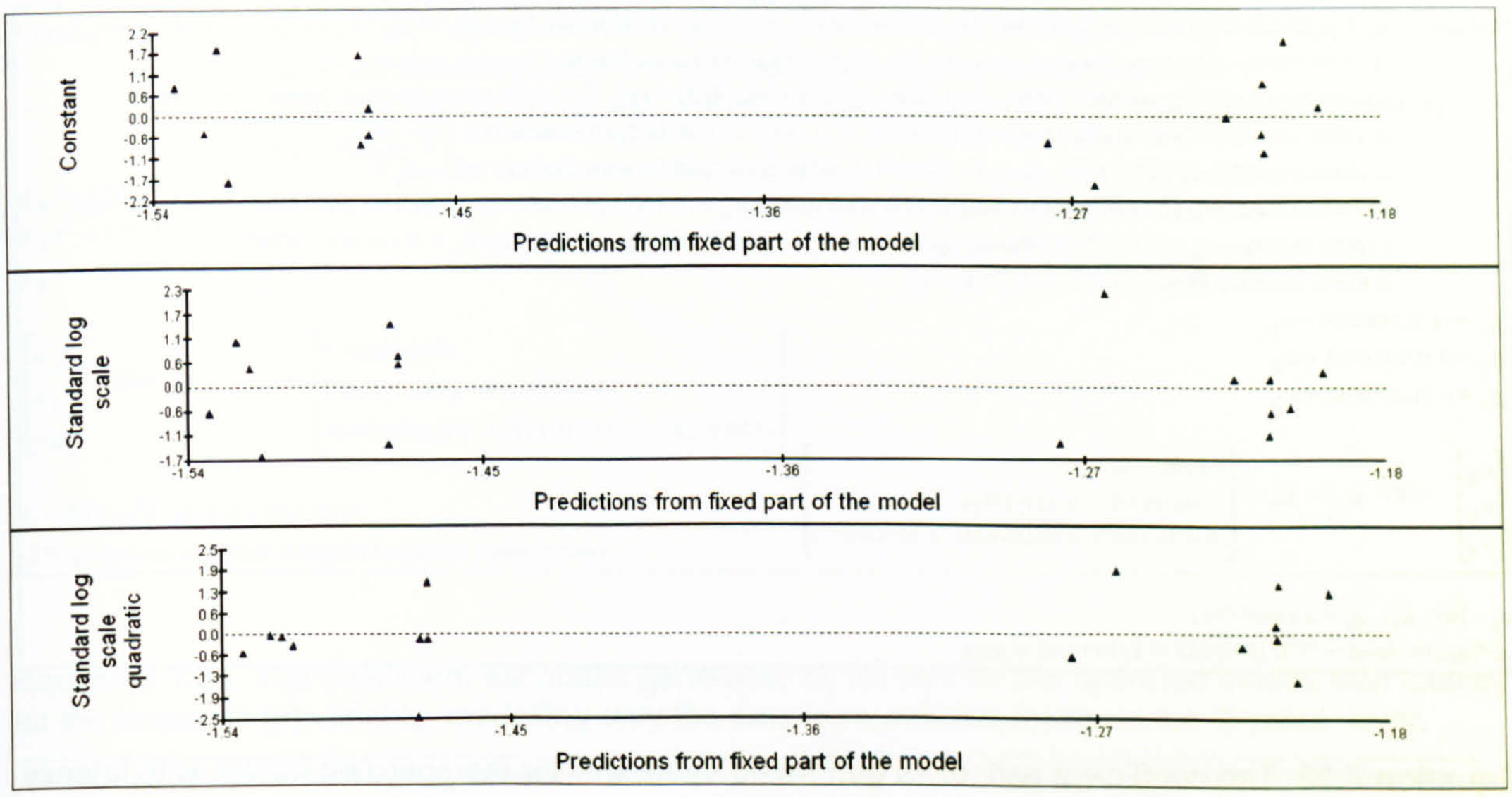

Figure 7.22 The standardised residuals (y-axis) plotted against the predicted values from the fixed part of the model (prior to backtransformation) for the Level 2 (i.e. *subject*) variance, from the model specified in Equation 7.34.

$$\begin{aligned}
\text{Latency}_{ij} = {} & \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -0.309(0.147)\text{Standard log scale cubic}_{ij} + 0.102(0.023)\text{2pell lvr press}_{ij} + \\
& -1.108(0.082)\text{2pell lvr press.Standard log scale}_{ij} + 0.054(0.100)\text{2pell lvr press.Standard log scale quadratic}_{ij} + \\
& 1.125(0.209)\text{2pell lvr press.Standard log scale cubic}_{ij} + 0.090(0.081)\text{4kHz=2pell}_{j} + -0.247(0.096)\text{4kHz=2pell.Standard log scale}_{ij} + \\
& -0.159(0.110)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + -0.006(0.205)\text{4kHz=2pell.Standard log scale cubic}_{ij} + \\
& -0.126(0.030)\text{4kHz=2pell.2pell lvr press}_{ij} + 0.639(0.110)\text{4kHz=2pell.2pell lvr press.Standard log scale}_{ij} + \\
& 0.209(0.137)\text{4kHz=2pell.2pell lvr press.Standard log scale quadratic}_{ij} + -0.573(0.285)\text{4kHz=2pell.2pell lvr press.Standard log scale cubic}_{ij} + \\
& 0.101(0.017)\text{Phase2}_{ij} + -0.042(0.018)\text{Phase2.2pell lvr press}_{ij} + -0.048(0.018)\text{Phase2.4kHz=2pell}_{ij} + -0.212(0.073)\text{UHT}_{j} + \\
& -0.045(0.017)\text{UHT.Phase2}_{ij} + e_{ij}
\end{aligned}$$

$$\beta_{0j} = -1.372(0.067) + u_{0j}$$
$$\beta_{1j} = 0.390(0.072) + u_{1j}$$
$$\beta_{2j} = 0.257(0.067) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.024(0.009) & & \\ -0.001(0.004) & 0.009(0.004) & \\ -0.007(0.006) & 0.004(0.004) & 0.013(0.008) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.155(0.002)$$

*-2*loglikelihood* = 8057.236(8164 of 8164 cases in use)

Equation 7.35 The coefficient estimates generated by MLwiN for the specified model, with latency as the response (*y*) variable, modelling only those (untransformed) latency datapoints under or equal to 3 seconds.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + \text{-0 245(0 153)Standard log scale cubic}_{ij} + \text{0 135(0.024)2pell lvr press}_{ij} +$$
$$\text{-1 247(0.084)2pell lvr press.Standard log scale}_{ij} + \text{0 135(0 098)2pell lvr press.Standard log scale quadratic}_{ij} +$$
$$\text{1 130(0.212)2pell lvr press.Standard log scale cubic}_{ij} + \text{0 121(0 088)4kHz=2pell}_{j} + \text{-0 202(0 101)4kHz=2pell Standard log scale}_{ij} +$$
$$\text{-0 218(0 111)4kHz=2pell Standard log scale quadratic}_{ij} + \text{-0 213(0 213)4kHz=2pell Standard log scale cubic}_{ij} +$$
$$\text{-0.181(0 032)4kHz=2pell 2pell lvr press}_{ij} + \text{0 649(0 115)4kHz=2pell 2pell lvr press Standard log scale}_{ij} +$$
$$\text{0.229(0.138)4kHz=2pell.2pell lvr press.Standard log scale quadratic}_{ij} + \text{-0 379(0 293)4kHz=2pell 2pell lvr press Standard log scale cubic}_{ij} +$$
$$\text{0 100(0 018)Phase2}_{ij} + \text{-0 049(0 019)Phase2.2pell lvr press}_{ij} + \text{-0 067(0 018)Phase2.4kHz=2pell}_{ij} + \text{-0 235(0 077)UHT}_{j} +$$
$$\text{-0 055(0 018)UHT.Phase2}_{ij} + \text{-0 019(0.077)R}_{j} + e_{ij}$$

$$\beta_{0j} = \text{-1 321(0 083)} + u_{0j}$$
$$\beta_{1j} = \text{0 396(0 075)} + u_{1j}$$
$$\beta_{2j} = \text{0 312(0 066)} + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} \text{0 028(0 010)} & & \\ \text{-0.003(0.005)} & \text{0 011(0.004)} & \\ \text{-0 009(0.007)} & \text{0 004(0 004)} & \text{0 010(0 007)} \end{bmatrix}$$

$$e_{ij} \sim N(0, \sigma_e^2) \quad \sigma_e^2 = \text{0 180(0 003)}$$

*-2*loglikelihood* = 9772.140(8619 of 8619 cases in use)

Equation 7.36 The coefficient estimates generated by MLwiN for the specified model, with latency as the response (*y*) variable.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + \text{0.940(0.142)Standard log scale cubic}_{ij} +$$
$$\text{0.452(0.102)4kHz=2pell.Standard log scale}_{ij} + \text{0.054(0.138)4kHz=2pell.Standard log scale quadratic}_{ij} +$$
$$\text{-0 671(0.194)4kHz=2pell.Standard log scale cubic}_{ij} + \text{-0.096(0.093)4kHz=2pell}_{j} + \text{-0.014(0.012)Phase2}_{ij} + e_{ij}$$

$$\beta_{0j} = \text{-1.291(0.067)} + u_{0j}$$
$$\beta_{1j} = \text{-0.847(0.072)} + u_{1j}$$
$$\beta_{2j} = \text{0.405(0.107)} + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} \text{0.033(0.012)} & & \\ \text{0.005(0.008)} & \text{0.020(0.009)} & \\ \text{0.003(0.012)} & \text{-0.011(0.011)} & \text{0.040(0.022)} \end{bmatrix}$$

$$e_{ij} \sim N(0, \sigma_e^2) \quad \sigma_e^2 = \text{0.166(0.003)}$$

*-2*loglikelihood* = 4907.670(4594 of 4594 cases in use)

Equation 7.37 The coefficient estimates generated by MLwiN for the specified model, with latency as the response (*y*) variable, modelling only the data from presses made on the '2-pellet' lever.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + \text{0 936(0 142)Standard log scale cubic}_{ij} +$$
$$\text{0 450(0.102)4kHz=2pell.Standard log scale}_{ij} + \text{0 052(0.138)4kHz=2pell.Standard log scale quadratic}_{ij} +$$
$$\text{-0.665(0.194)4kHz=2pell.Standard log scale cubic}_{ij} + \text{-0.095(0 079)4kHz=2pell}_{j} + \text{-0.014(0 012)Phase2}_{ij} + \text{-0.201(0.078)UHT}_{j} + e_{ij}$$

$$\beta_{0j} = \text{-1.191(0 069)} + u_{0j}$$
$$\beta_{1j} = \text{-0.846(0 072)} + u_{1j}$$
$$\beta_{2j} = \text{0.407(0.107)} + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} \text{0 023(0.009)} & & \\ \text{0.003(0.006)} & \text{0.021(0.009)} & \\ \text{-0.002(0.010)} & \text{-0.010(0 011)} & \text{0.040(0 022)} \end{bmatrix}$$

$$e_{ij} \sim N(0, \sigma_e^2) \quad \sigma_e^2 = \text{0.166(0.003)}$$

*-2*loglikelihood* = 4902.605(4594 of 4594 cases in use)

Equation 7.38 The coefficient estimates generated by MLwiN for the specified model, with latency as the response (*y*) variable, modelling only the data from presses made on the '2-pellet' lever.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + 0\,951(0.142)\text{Standard log scale cubic}_{ij} +$$
$$0.462(0.103)4\text{kHz=2pell.Standard log scale}_{ij} + 0.055(0.139)4\text{kHz=2pell.Standard log scale quadratic}_{ij} +$$
$$-0.689(0\,194)4\text{kHz=2pell.Standard log scale cubic}_{ij} + -0\,098(0\,079)4\text{kHz=2pell}_{j} + 0.029(0\,017)\text{Phase2}_{ij} + -0\,157(0\,079)\text{UHT}_{j} +$$
$$-0\,086(0\,024)\text{UHT.Phase2}_{ij} + e_{ij}$$

$$\beta_{0j} = -1.210(0\,069) + u_{0j}$$

$$\beta_{1j} = -0\,852(0\,072) + u_{1j}$$

$$\beta_{2j} = 0.400(0.107) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0\,024(0\,009) & & \\ 0.003(0.006) & 0.021(0.009) & \\ -0.002(0\,010) & -0\,011(0\,011) & 0\,041(0.022) \end{bmatrix}$$

$$e_{ij} \sim N(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.166(0\,003)$$

$$-2*\textit{loglikelihood} = 4889.834(4594 \text{ of } 4594 \text{ cases in use})$$

Equation 7.39 The coefficient estimates generated by MLwiN for the specified model, with latency as the response (y) variable, modelling only the data from presses made on the '2-pellet' lever.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + 0.830(0.193)\text{Standard log scale cubic}_{ij} + -0\,102(0.079)4\text{kHz=2pell}_{j} +$$
$$0\,473(0\,102)4\text{kHz=2pell.Standard log scale\#2}_{ij} + 0\,064(0.136)4\text{kHz=2pell.Standard log scale quadratic\#2}_{ij} +$$
$$-0.722(0.194)4\text{kHz=2pell Standard log scale cubic\#2}_{ij} + 0\,054(0\,026)\text{Phase2}_{ij} + -0\,118(0\,092)\text{Phase2.Standard log scale\#3}_{ij} +$$
$$-0.134(0\,106)\text{Phase2.Standard log scale quadratic\#3}_{ij} + 0\,391(0.222)\text{Phase2.Standard log scale cubic\#3}_{ij} + -0\,121(0\,081)\text{UHT}_{j} +$$
$$-0.155(0\,037)\text{Phase2.UHT}_{ij} + -0.217(0.119)\text{UHT.Standard log scale\#1}_{ij} + -0.188(0\,144)\text{UHT.Standard log scale quadratic\#1}_{ij} +$$
$$0.307(0.235)\text{UHT.Standard log scale cubic\#1}_{ij} + 0\,430(0.131)\text{Phase2.UHT.Standard log scale\#1}_{ij} +$$
$$0\,158(0\,149)\text{Phase2.UHT Standard log scale quadratic\#1}_{ij} + -0.817(0\,313)\text{Phase2.UHT.Standard log scale cubic\#1}_{ij} + e_{ij}$$

$$\beta_{0j} = -1\,221(0\,070) + u_{0j}$$

$$\beta_{1j} = -0.798(0\,096) + u_{1j}$$

$$\beta_{2j} = 0\,509(0\,130) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.024(0.009) & & \\ 0.003(0\,006) & 0\,021(0\,009) & \\ -0\,001(0\,010) & -0\,013(0\,011) & 0\,037(0\,021) \end{bmatrix}$$

$$e_{ij} \sim N(0,\ \sigma_e^2) \quad \sigma_e^2 = 0\,165(0.003)$$

$$-2*\textit{loglikelihood} = 4869.060(4594 \text{ of } 4594 \text{ cases in use})$$

Equation 7.40 The coefficient estimates generated by MLwiN for the specified model, with latency as the response (y) variable, modelling only the data from presses made on the '2-pellet' lever.

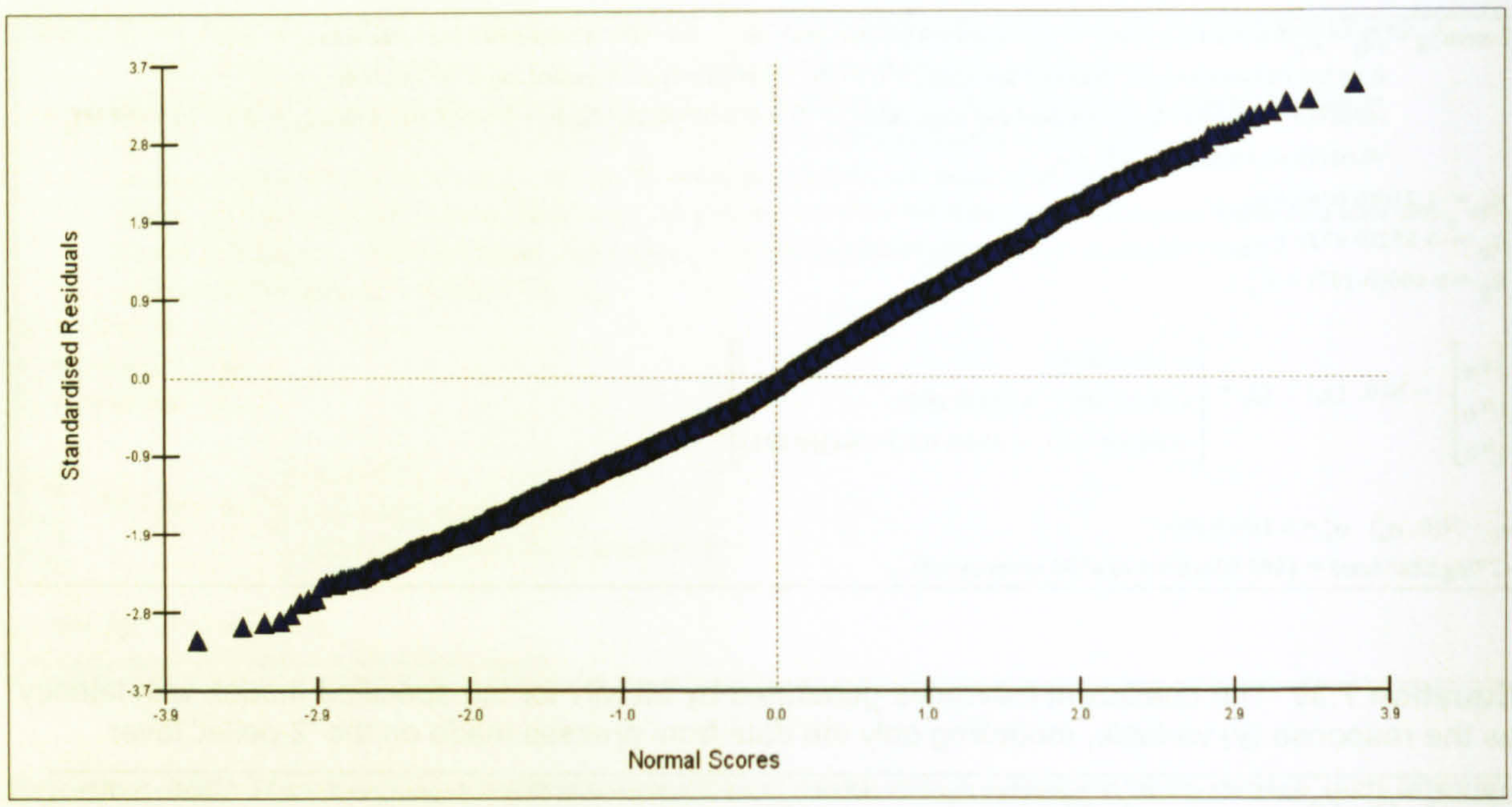

Figure 7.23 The standardised residuals (y-axis) plotted against the Normal scores for the Level 1 (i.e. *trial*) variance, from the model specified in Equation 7.40, modelling only the data from presses made on the '2-pellet' lever.

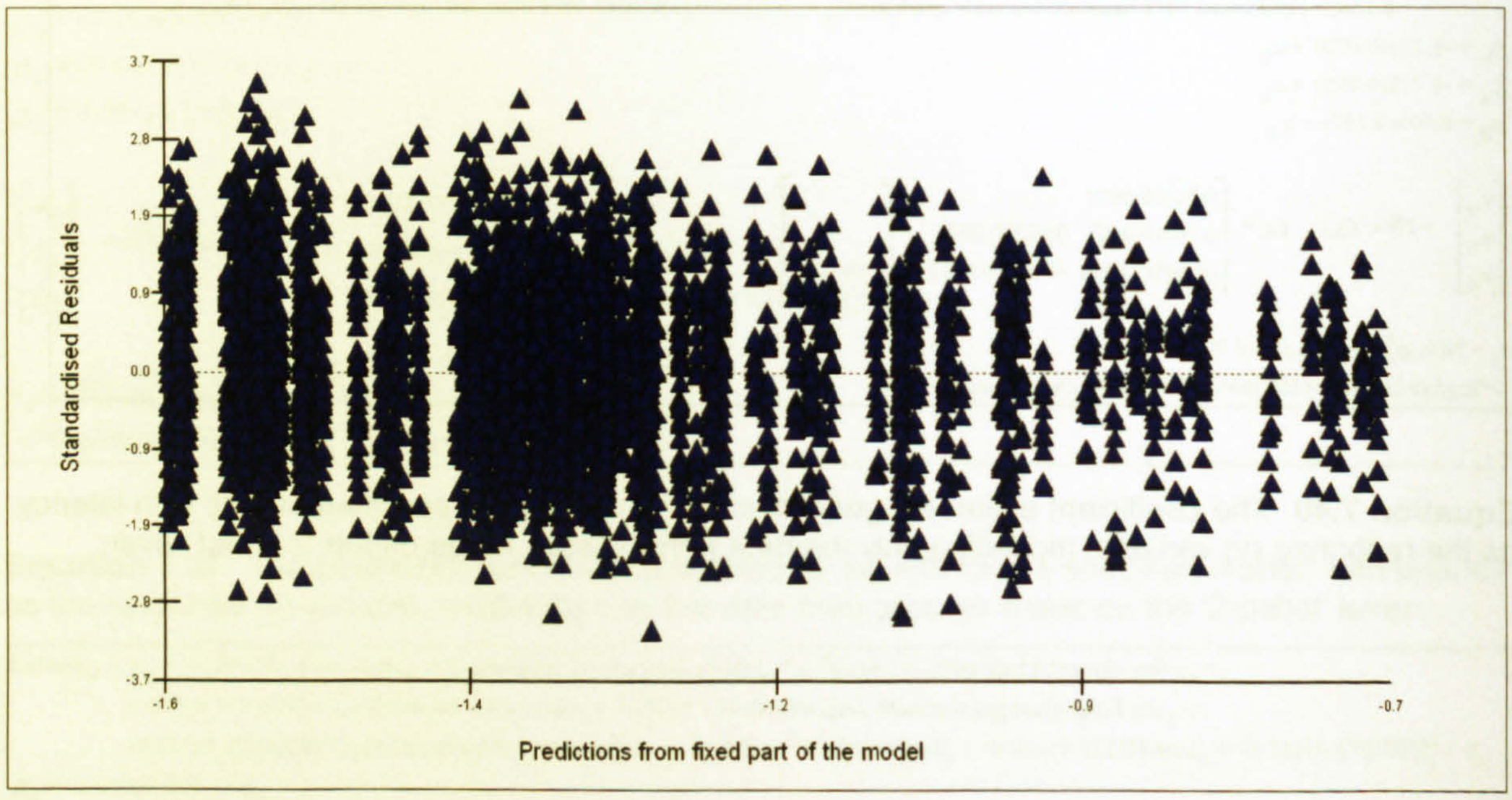

Figure 7.24 The standardised residuals (y-axis) plotted against the predicted values from the fixed part of the model (prior to backtransformation) for the Level 1 (i.e. *trial*) variance, from the model specified in Equation 7.40, modelling only the data from presses made on the '2-pellet' lever.

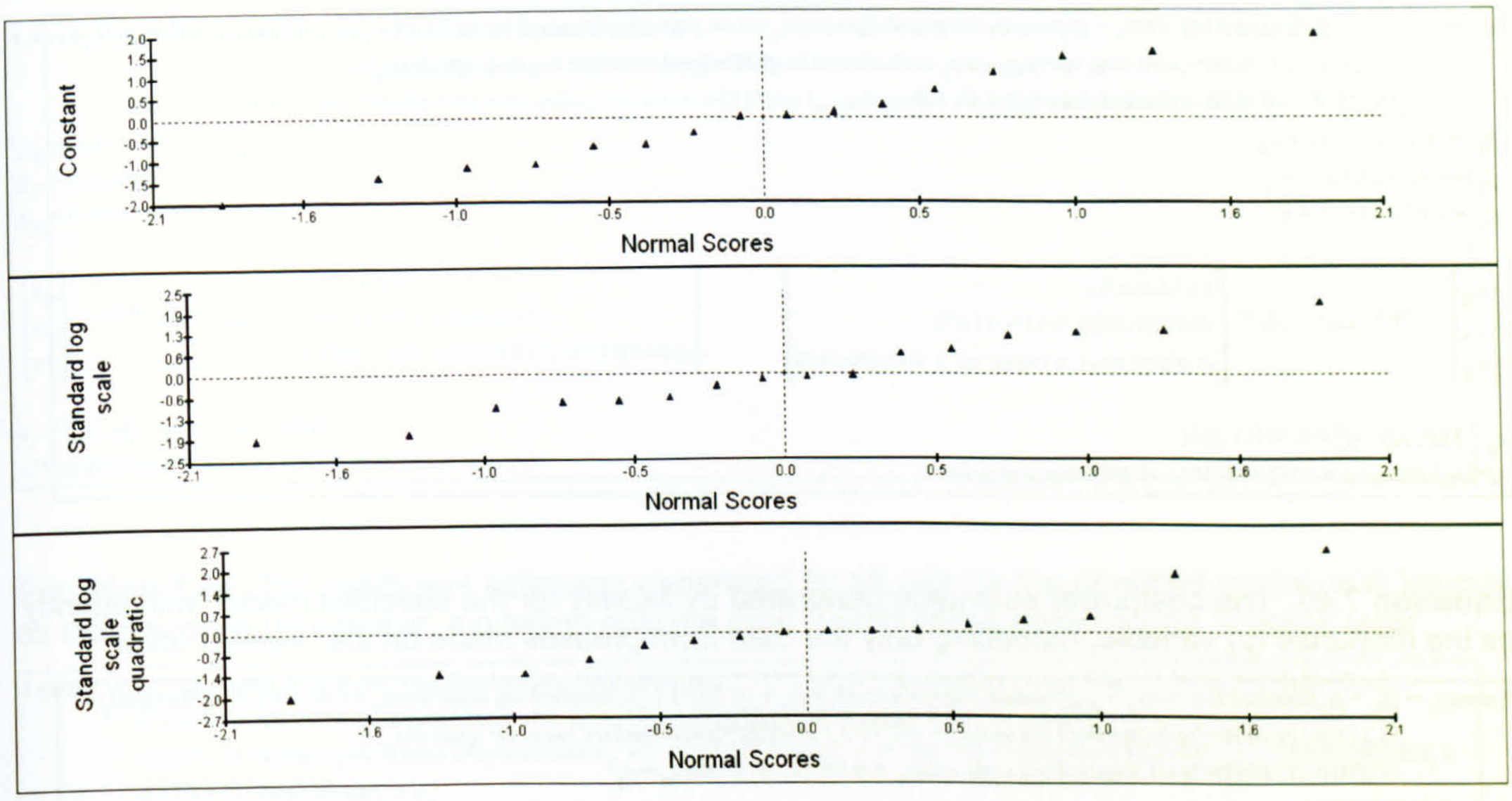

Figure 7.25 The standardised residuals (y-axis) plotted against the Normal scores for the Level 2 (i.e. *subject*) variance, from the model specified in Equation 7.40, modelling only the data from presses made on the '2-pellet' lever.

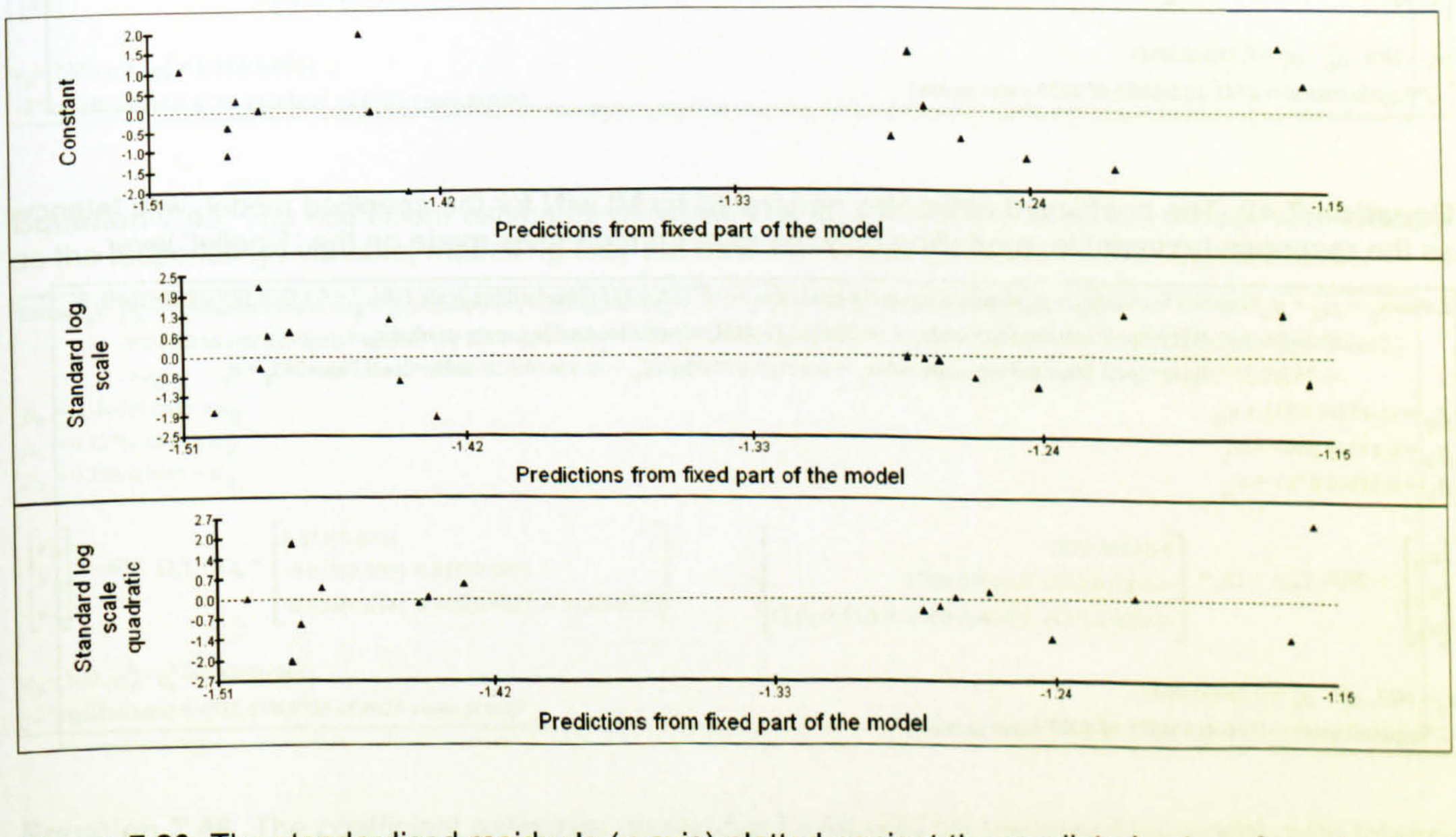

Figure 7.26 The standardised residuals (y-axis) plotted against the predicted values from the fixed part of the model (prior to backtransformation) for the Level 2 (i.e. *subject*) variance, from the model specified in Equation 7.40, modelling only the data from presses made on the '2-pellet' lever.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + \text{-0 234(0 156)Standard log scale cubic}_{ij} + 0.082(0\ 118)\text{4kHz=2pell}_{j} + \text{-0.216(0 107)4kHz=2pell Standard log scale}_{ij} + \text{-0 256(0.115)4kHz=2pell.Standard log scale quadratic}_{ij} + \text{-0.251(0.218)4kHz=2pell.Standard log scale cubic}_{ij} + e_{ij}$$

$$\beta_{0j} = \text{-1 404(0.083)} + u_{0j}$$

$$\beta_{1j} = 0\ 423(0.080) + u_{1j}$$

$$\beta_{2j} = 0.350(0.069) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0\ 054(0.020) & & \\ \text{-0.011(0 009)} & 0\ 015(0.007) & \\ \text{-0 028(0 014)} & 0\ 006(0\ 007) & 0\ 012(0\ 013) \end{bmatrix}$$

$$e_{ij} \sim N(0,\ \sigma_e^2) \quad \sigma_e^2 = 0\ 187(0\ 004)$$

$$-2*\textit{loglikelihood} = 4752.508(4025 \text{ of } 4025 \text{ cases in use})$$

Equation 7.41 The coefficient estimates generated by MLwiN for the specified model, with latency as the response (y) variable, modelling only the data from presses made on the '1-pellet' lever.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + \text{-0.229(0 156)Standard log scale cubic}_{ij} + 0\ 084(0\ 118)\text{4kHz=2pell}_{j} + \text{-0.212(0.108)4kHz=2pell Standard log scale}_{ij} + \text{-0.256(0.115)4kHz=2pell.Standard log scale quadratic}_{ij} + \text{-0 254(0 217)4kHz=2pell Standard log scale cubic}_{ij} + 0\ 046(0\ 014)\text{Phase2}_{ij} + e_{ij}$$

$$\beta_{0j} = \text{-1 429(0 084)} + u_{0j}$$

$$\beta_{1j} = 0\ 419(0\ 080) + u_{1j}$$

$$\beta_{2j} = 0.349(0\ 069) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0\ 053(0.020) & & \\ \text{-0.011(0 009)} & 0.015(0\ 007) & \\ \text{-0 028(0.014)} & 0.006(0.007) & 0\ 012(0.013) \end{bmatrix}$$

$$e_{ij} \sim N(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.186(0\ 004)$$

$$-2*\textit{loglikelihood} = 4741\ 132(4025 \text{ of } 4025 \text{ cases in use})$$

Equation 7.42 The coefficient estimates generated by MLwiN for the specified model, with latency as the response (y) variable, modelling only the data from presses made on the '1-pellet' lever.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + \text{-0.222(0 155)Standard log scale cubic}_{ij} + 0.138(0\ 119)\text{4kHz=2pell}_{j} + \text{-0 207(0.107)4kHz=2pell Standard log scale}_{ij} + \text{-0.254(0.115)4kHz=2pell.Standard log scale quadratic}_{ij} + \text{-0 263(0.217)4kHz=2pell Standard log scale cubic}_{ij} + 0\ 091(0\ 018)\text{Phase2}_{ij} + \text{-0.106(0 028)4kHz=2pell Phase2\#2}_{ij} + e_{ij}$$

$$\beta_{0j} = \text{-1 454(0 084)} + u_{0j}$$

$$\beta_{1j} = 0.415(0\ 080) + u_{1j}$$

$$\beta_{2j} = 0.348(0.070) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0\ 053(0.019) & & \\ \text{-0.012(0 009)} & 0.015(0\ 007) & \\ \text{-0.028(0 014)} & 0.006(0.007) & 0.013(0.013) \end{bmatrix}$$

$$e_{ij} \sim N(0,\ \sigma_e^2) \quad \sigma_e^2 = 0\ 186(0\ 004)$$

$$-2*\textit{loglikelihood} = 4726.413(4025 \text{ of } 4025 \text{ cases in use})$$

Equation 7.43 The coefficient estimates generated by MLwiN for the specified model, with latency as the response (y) variable, modelling only the data from presses made on the '1-pellet' lever.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + \text{-0.221(0 155)Standard log scale cubic}_{ij} + \text{0 139(0.102)4kHz=2pell}_{j} + \text{-0 204(0.106)4kHz=2pell.Standard log scale}_{ij} + \text{-0 251(0 114)4kHz=2pell.Standard log scale quadratic}_{ij} + \text{-0 263(0 217)4kHz=2pell Standard log scale cubic}_{ij} + \text{0 091(0 018)Phase2}_{ij} + \text{-0.107(0 028)4kHz=2pell Phase2\#2}_{ij} + \text{-0 201(0 066)UHT}_{j} + e_{ij}$$

$$\beta_{0j} = \text{-1.354(0 079)} + u_{0j}$$

$$\beta_{1j} = \text{0.413(0 079)} + u_{1j}$$

$$\beta_{2j} = \text{0 350(0 068)} + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} \text{0 039(0 014)} & & \\ \text{-0 007(0 007)} & \text{0 014(0 007)} & \\ \text{-0 024(0.012)} & \text{0 006(0 007)} & \text{0 012(0 013)} \end{bmatrix}$$

$$e_{ij} \sim N(0, \sigma_e^2) \quad \sigma_e^2 = \text{0 186(0.004)}$$

$$\textit{-2*loglikelihood} = \text{4719 563(4025 of 4025 cases in use)}$$

Equation 7.44 The coefficient estimates generated by MLwiN for the specified model, with latency as the response (*y*) variable, modelling only the data from presses made on the '1-pellet' lever.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + \text{-0 221(0 155)Standard log scale cubic}_{ij} + \text{0.140(0 102)4kHz=2pell}_{j} + \text{-0 206(0.106)4kHz=2pell Standard log scale}_{ij} + \text{-0 252(0 114)4kHz=2pell Standard log scale quadratic}_{ij} + \text{-0 260(0.217)4kHz=2pell.Standard log scale cubic}_{ij} + \text{0 107(0 023)Phase2}_{ij} + \text{-0 108(0.028)4kHz=2pell Phase2\#2}_{ij} + \text{-0 186(0 068)UHT}_{j} + \text{-0 029(0.027)UHT.Phase2}_{ij} + e_{ij}$$

$$\beta_{0j} = \text{-1 363(0 079)} + u_{0j}$$

$$\beta_{1j} = \text{0 413(0 079)} + u_{1j}$$

$$\beta_{2j} = \text{0.352(0 068)} + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} \text{0 039(0 014)} & & \\ \text{-0 007(0 007)} & \text{0 014(0 007)} & \\ \text{-0 025(0 012)} & \text{0 006(0 007)} & \text{0 012(0 013)} \end{bmatrix}$$

$$e_{ij} \sim N(0, \sigma_e^2) \quad \sigma_e^2 = \text{0 186(0 004)}$$

$$\textit{-2*loglikelihood} = \text{4718.442(4025 of 4025 cases in use)}$$

Equation 7.45 The coefficient estimates generated by MLwiN for the specified model, with latency as the response (*y*) variable, modelling only the data from presses made on the '1-pellet' lever.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + \text{-0 279(0 073)4kHz=2pell Standard log scale}_{ij} + \text{-0 044(0 099)4kHz=2pell Standard log scale quadratic}_{ij} + \text{0 113(0 118)4kHz=2pell}_{j} + \text{0 091(0 018)Phase2}_{ij} + \text{-0 106(0 028)4kHz=2pell Phase2}_{ij} + e_{ij}$$

$$\beta_{0j} = \text{-1 461(0.084)} + u_{0j}$$

$$\beta_{1j} = \text{0.327(0.051)} + u_{1j}$$

$$\beta_{2j} = \text{0 330(0 069)} + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} \text{0 053(0 020)} & & \\ \text{-0 012(0 009)} & \text{0 016(0 007)} & \\ \text{-0.028(0 014)} & \text{0 008(0 008)} & \text{0 014(0 014)} \end{bmatrix}$$

$$e_{ij} \sim N(0, \sigma_e^2) \quad \sigma_e^2 = \text{0 186(0 004)}$$

$$\textit{-2*loglikelihood} = \text{4738 659(4025 of 4025 cases in use)}$$

Equation 7.46 The coefficient estimates generated by MLwiN for the specified model, with latency as the response (*y*) variable, modelling only the data from presses made on the '1-pellet' lever.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + \text{-0.281(0.067)4kHz=2pell Standard log scale}_{ij} +$$
$$\text{-0.048(0.100)4kHz=2pell Standard log scale quadratic}_{ij} + \text{0 113(0 115)4kHz=2pell}_{j} + \text{0.121(0 031)Phase2}_{ij} +$$
$$\text{-0 101(0.051)Phase2 Standard log scale}_{ij} + \text{-0.169(0.105)Phase2.Standau d log scale quadratic}_{ij} + \text{-0 046(0.040)UHT.Phase2}_{ij} +$$
$$\text{0.006(0 075)UHT.Standard log scale}_{ij} + \text{-0 164(0.091)UHT.Standard log scale quadratic}_{ij} + \text{-0.103(0 028)4kHz=2pell Phase2}_{ij} +$$
$$\text{0.068(0.073)UHT.Phase2.Standard log scale}_{ij} + \text{0.118(0.148)UHT.Phase2.Standard log scale quadratic}_{ij} + e_{ij}$$

$$\beta_{0j} = \text{-1.465(0 082)} + u_{0j}$$
$$\beta_{1j} = \text{0.361(0 063)} + u_{1j}$$
$$\beta_{2j} = \text{0.476(0 093)} + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) \; : \; \Omega_u = \begin{bmatrix} \text{0.050(0.018)} & & \\ \text{-0 009(0 008)} & \text{0 012(0 006)} & \\ \text{-0 033(0 015)} & \text{0 006(0 007)} & \text{0.016(0.014)} \end{bmatrix}$$

$$e_{ij} \sim N(0, \sigma_e^2) \quad \sigma_e^2 = \text{0 186(0.004)}$$
$$\text{-2*loglikelihood} = \text{4728.050(4025 of 4025 cases in use)}$$

Equation 7.47 The coefficient estimates generated by MLwiN for the specified model, with latency as the response (*y*) variable, modelling only the data from presses made on the '1-pellet' lever.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + \text{-0 221(0 155)Standard log scale cubic}_{ij} + \text{0.139(0.102)4kHz=2pell}_{j} +$$
$$\text{-0 204(0.106)4kHz=2pell.Standau d log scale}_{ij} + \text{-0.251(0.113)4kHz=2pell.Standard log scale quadratic}_{ij} +$$
$$\text{-0 263(0.217)4kHz=2pell.Standau d log scale cubic}_{ij} + \text{0.091(0 018)Phase2}_{ij} + \text{-0 107(0.028)4kHz=2pell Phase2\#1}_{ij} + \text{-0 201(0.066)UHT}_{j} + e_{ij}$$

$$\beta_{0j} = \text{-1.354(0.079)} + u_{0j}$$
$$\beta_{1j} = \text{0.413(0 079)} + u_{1j}$$
$$\beta_{2j} = \text{0.350(0.068)} + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) \; : \; \Omega_u = \begin{bmatrix} \text{0 039(0.014)} & & \\ \text{-0 007(0 007)} & \text{0.014(0.007)} & \\ \text{-0 024(0 012)} & \text{0 006(0.007)} & \text{0.011(0.013)} \end{bmatrix}$$

$$e_{ij} \sim N(0, \sigma_e^2) \quad \sigma_e^2 = \text{0 186(0 004)}$$
$$\text{-2*loglikelihood} = \text{4719.562(4025 of 4025 cases in use)}$$

Equation 7.48 The coefficient estimates generated by MLwiN for the specified model, with latency as the response (*y*) variable, modelling only the data from presses made on the '1-pellet' lever.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + \text{-0 218(0 155)Standard log scale cubic}_{ij} + \text{0.120(0.123)4kHz=2pell}_{j} +$$
$$\text{-0.215(0 105)4kHz=2pell Standau d log scale}_{ij} + \text{-0.242(0.113)4kHz=2pell Standau d log scale quadratic}_{ij} +$$
$$\text{-0 247(0 216)4kHz=2pell Standard log scale cubic}_{ij} + \text{0 056(0.026)Phase2}_{ij} + \text{0.006(0.039)4kHz=2pell Phase2\#1}_{ij} + \text{-0.195(0 092)UHT}_{j} +$$
$$\text{0.068(0 036)Phase2.UHT}_{ij} + \text{0 032(0.131)4kHz=2pell.UHT}_{j} + \text{-0 228(0 055)4kHz=2pell Phase2 UHT}_{ij} + e_{ij}$$

$$\beta_{0j} = \text{-1.355(0.086)} + u_{0j}$$
$$\beta_{1j} = \text{0.413(0.078)} + u_{1j}$$
$$\beta_{2j} = \text{0.348(0.068)} + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) \; : \; \Omega_u = \begin{bmatrix} \text{0 040(0 015)} & & \\ \text{-0 007(0 007)} & \text{0.013(0 007)} & \\ \text{-0 026(0 012)} & \text{0 005(0 007)} & \text{0 012(0 013)} \end{bmatrix}$$

$$e_{ij} \sim N(0, \sigma_e^2) \quad \sigma_e^2 = \text{0 185(0 004)}$$
$$\text{-2*loglikelihood} = \text{4701.103(4025 of 4025 cases in use)}$$

Equation 7.49 The coefficient estimates generated by MLwiN for the specified model, with latency as the response (*y*) variable, modelling only the data from presses made on the '1-pellet' lever.

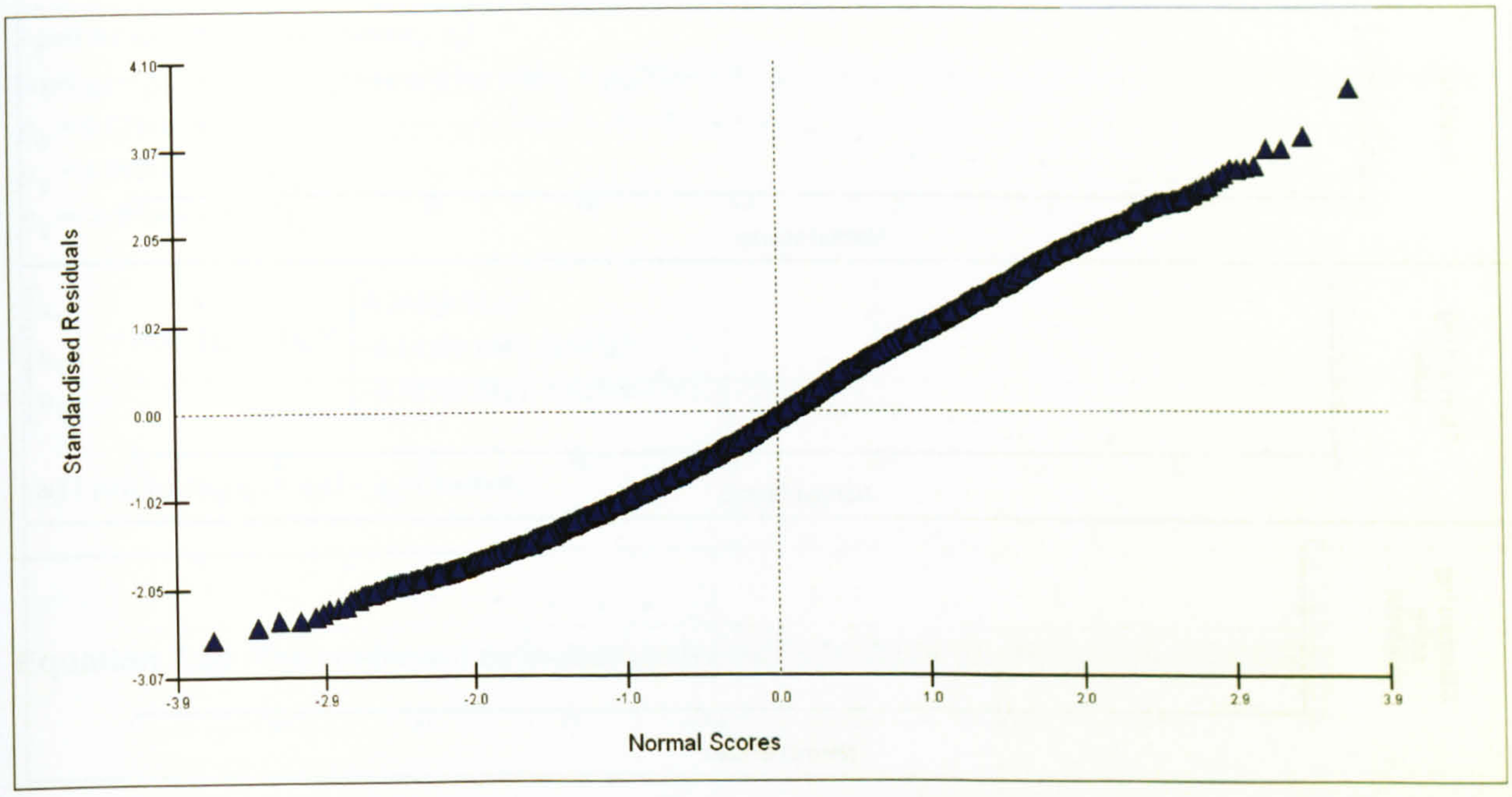

Figure 7.27 The standardised residuals (y-axis) plotted against the Normal scores for the Level 1 (i.e. *trial*) variance, from the model specified in Equation 7.49, modelling only the data from presses made on the '1-pellet' lever.

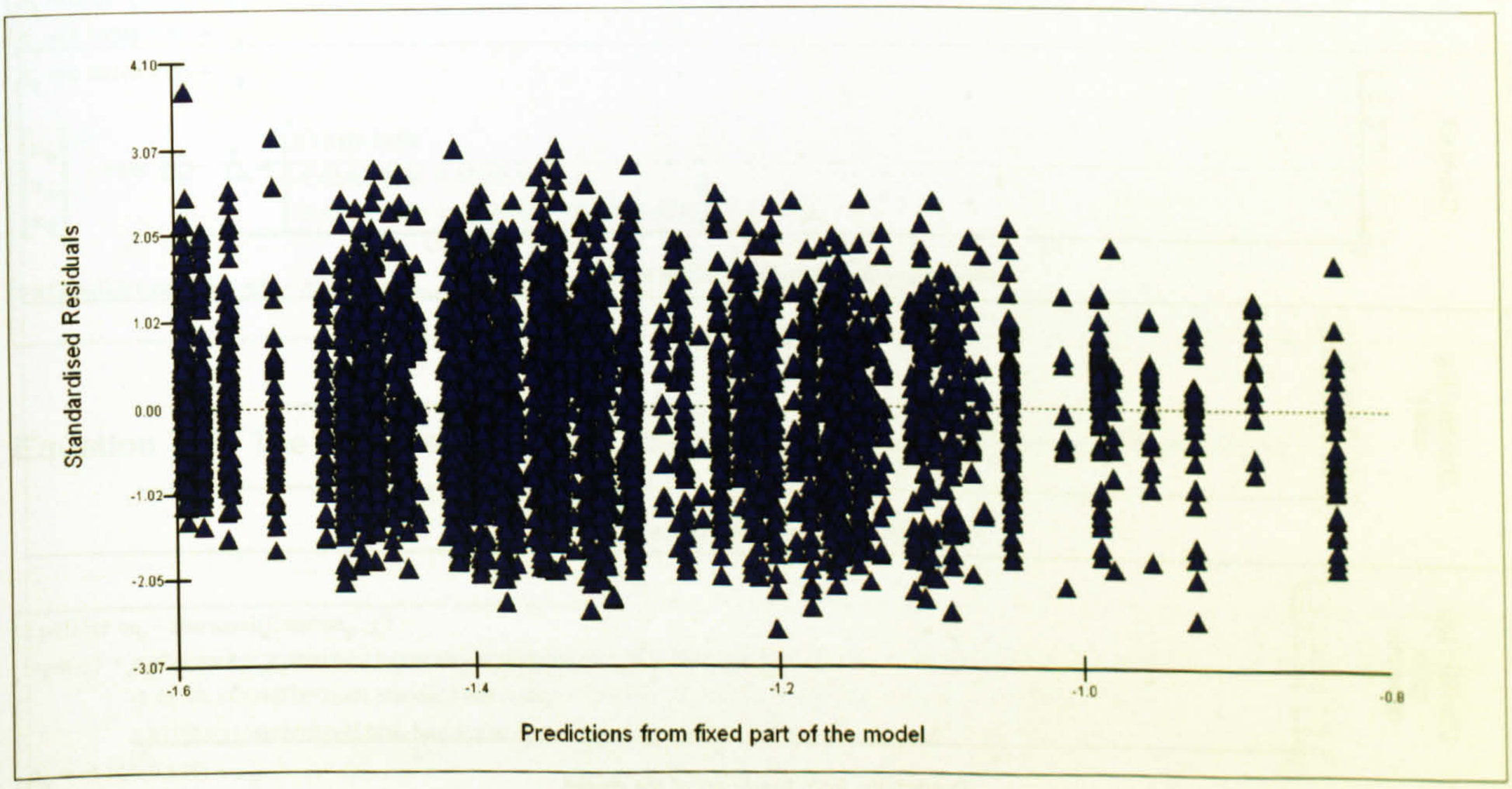

Figure 7.28 The standardised residuals (y-axis) plotted against the predicted values from the fixed part of the model (prior to backtransformation) for the Level 1 (i.e. *trial*) variance, from the model specified in Equation 7.49, modelling only the data from presses made on the '1-pellet' lever.

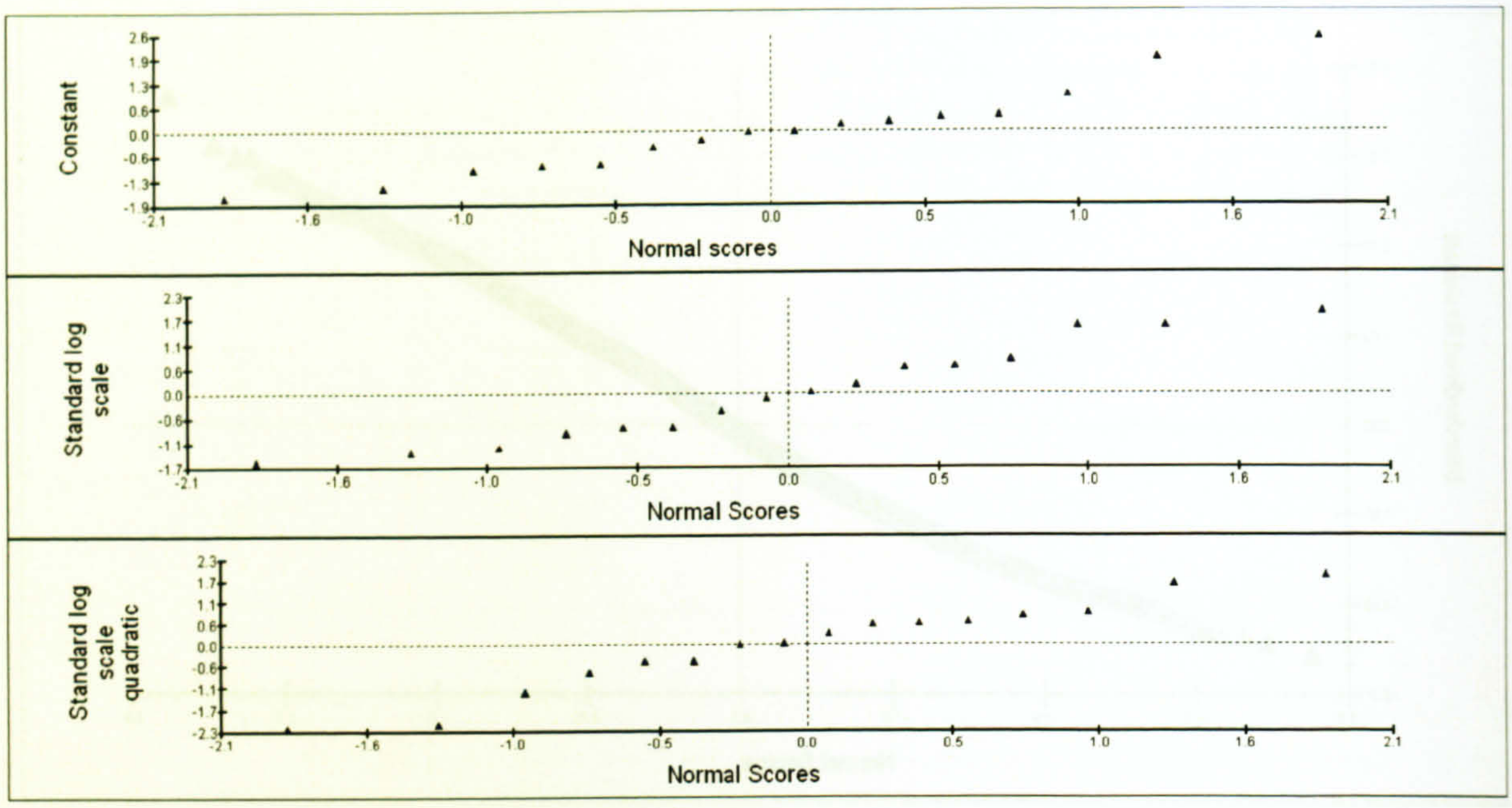

Figure 7.29 The standardised residuals (y-axis) plotted against the Normal scores for the Level 2 (i.e. *subject*) variance, from the model specified in Equation 7.49, modelling only the data from presses made on the '1-pellet' lever.

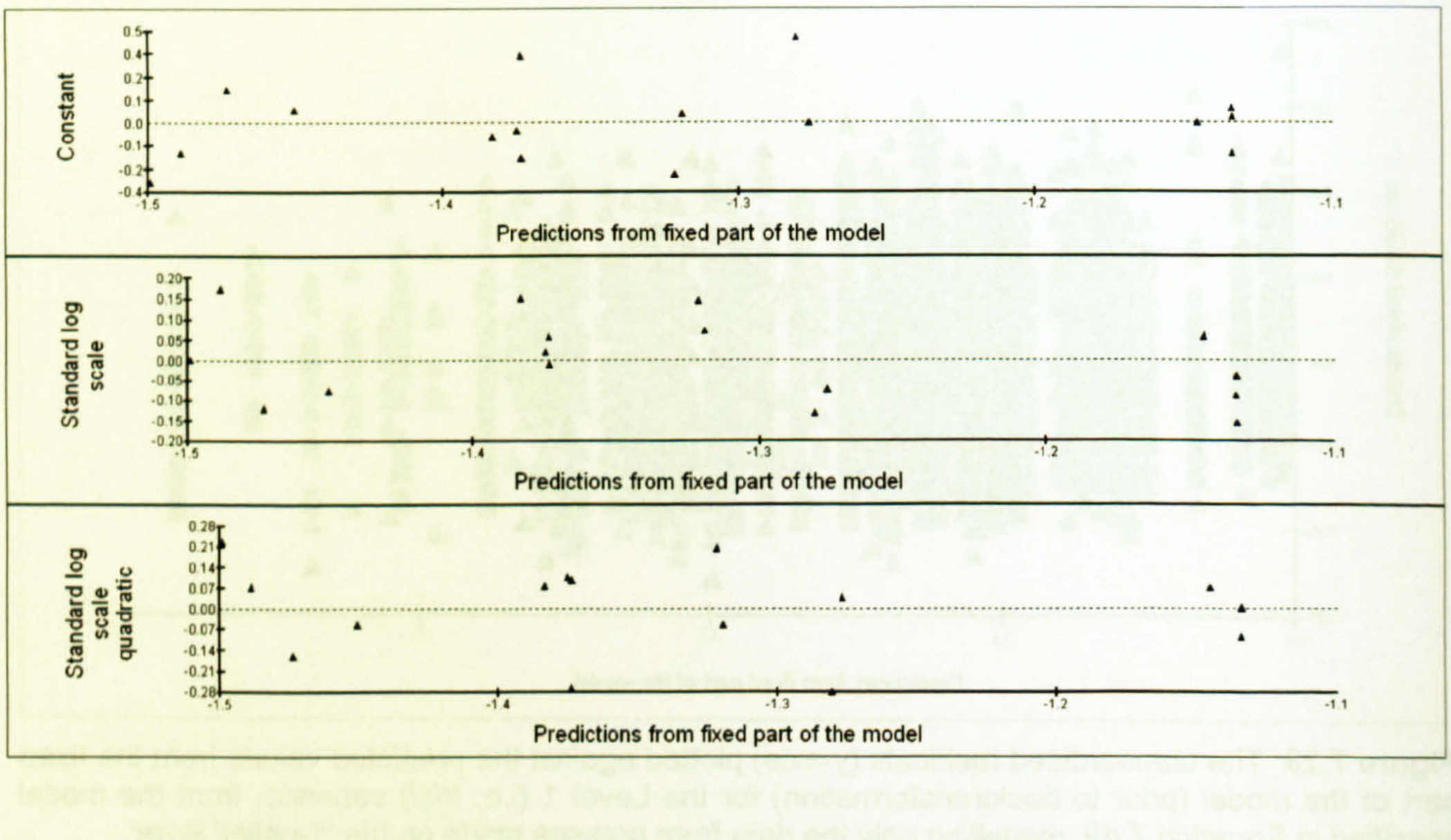

Figure 7.30 The standardised residuals (y-axis) plotted against the predicted values from the fixed part of the model (prior to backtransformation) for the Level 2 (i.e. *subject*) variance, from the model specified in Equation 7.49, modelling only the data from presses made on the '1-pellet' lever.

$$\text{2 pell lvr on}_{ij} \sim \text{Binomial}(\text{Constant}_{ij}, \pi_{ij})$$

$$\text{logit}(\pi_{ij}) = \beta_{0j}\text{Constant} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -6.339(0.296)\text{Standard log scale cubic}_{ij}$$

$$\beta_{0j} = 0.171(0.141) + u_{0j}$$

$$\beta_{1j} = 4.992(0.206) + u_{1j}$$

$$\beta_{2j} = -0.979(0.355) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.303(0.112) & & \\ -0.143(0.106) & 0.447(0.177) & \\ -0.507(0.242) & 0.619(0.294) & 1.798(0.714) \end{bmatrix}$$

$$\text{var}(\text{2 pell lvr on}_{ij} | \pi_{ij}) = \pi_{ij}(1 - \pi_{ij})/\text{Constant}_{ij}$$

Equation 7.50 The coefficient estimates generated by MLwiN for the specified model.

$$\text{2 pell lvr on}_{ij} \sim \text{Binomial}(\text{Constant}_{ij}, \pi_{ij})$$

$$\text{logit}(\pi_{ij}) = \beta_{0j}\text{Constant} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -7.780(0.460)\text{Standard log scale cubic}_{ij} + 0.855(0.186)\text{4kHz=2pell}_{j} + -1.436(0.388)\text{4kHz=2pell.Standard log scale}_{ij} + -2.160(0.526)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + 2.465(0.631)\text{4kHz=2pell.Standard log scale cubic}_{ij}$$

$$\beta_{0j} = -0.287(0.132) + u_{0j}$$

$$\beta_{1j} = 5.762(0.279) + u_{1j}$$

$$\beta_{2j} = 0.441(0.375) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.124(0.049) & & \\ -0.016(0.059) & 0.356(0.144) & \\ -0.079(0.097) & 0.346(0.179) & 0.766(0.348) \end{bmatrix}$$

$$\text{var}(\text{2 pell lvr on}_{ij} | \pi_{ij}) = \pi_{ij}(1 - \pi_{ij})/\text{Constant}_{ij}$$

Equation 7.51 The coefficient estimates generated by MLwiN for the specified model.

$$\text{2 pell lvr on}_{ij} \sim \text{Binomial}(\text{Constant}_{ij}, \pi_{ij})$$

$$\text{logit}(\pi_{ij}) = \beta_{0j}\text{Constant} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -7.795(0.461)\text{Standard log scale cubic}_{ij} + 0.856(0.186)\text{4kHz=2pell}_{j} + -1.437(0.389)\text{4kHz=2pell.Standard log scale}_{ij} + -2.164(0.526)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + 2.471(0.631)\text{4kHz=2pell.Standard log scale cubic}_{ij} + -0.170(0.053)\text{Phase\_2}_{ij} + -0.209(0.053)\text{Phase\_3}_{ij}$$

$$\beta_{0j} = -0.160(0.135) + u_{0j}$$

$$\beta_{1j} = 5.771(0.280) + u_{1j}$$

$$\beta_{2j} = 0.442(0.376) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.124(0.049) & & \\ -0.017(0.059) & 0.358(0.144) & \\ -0.080(0.097) & 0.347(0.180) & 0.769(0.349) \end{bmatrix}$$

$$\text{var}(\text{2 pell lvr on}_{ij} | \pi_{ij}) = \pi_{ij}(1 - \pi_{ij})/\text{Constant}_{ij}$$

Equation 7.52 The coefficient estimates generated by MLwiN for the specified model.

2 pell lvr on$_{ij}$ ~ Binomial(Constant$_{ij}$, π_{ij})

logit(π_{ij}) = β_{0j}Constant + β_{1j}Standard log scale$_{ij}$ + β_{2j}Standard log scale quadratic$_{ij}$ + -7.828(0.462)Standard log scale cubic$_{ij}$ + 0.655(0.196)4kHz=2pell$_j$ + -1.465(0.390)4kHz=2pell.Standard log scale$_{ij}$ + -2.164(0.527)4kHz=2pell.Standard log scale quadratic$_{ij}$ + 2.508(0.632)4kHz=2pell.Standard log scale cubic$_{ij}$ + -0.393(0.077)Phase_2$_{ij}$ + -0.303(0.077)Phase_3$_{ij}$ + 0.423(0.106)Phase_2.4kHz=2pell$_{ij}$ + 0.180(0.105)Phase_3.4kHz=2pell$_{ij}$

β_{0j} = -0.055(0.139) + u_{0j}

β_{1j} = 5.795(0.281) + u_{1j}

β_{2j} = 0.443(0.377) + u_{2j}

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) \; : \; \Omega_u = \begin{bmatrix} 0.124(0.049) & & \\ -0.017(0.060) & 0.359(0.145) & \\ -0.081(0.097) & 0.347(0.181) & 0.773(0.351) \end{bmatrix}$$

var(2 pell lvr on$_{ij}$|π_{ij}) = $\pi_{ij}(1 - \pi_{ij})$/Constant$_{ij}$

Equation 7.53 The coefficient estimates generated by MLwiN for the specified model.

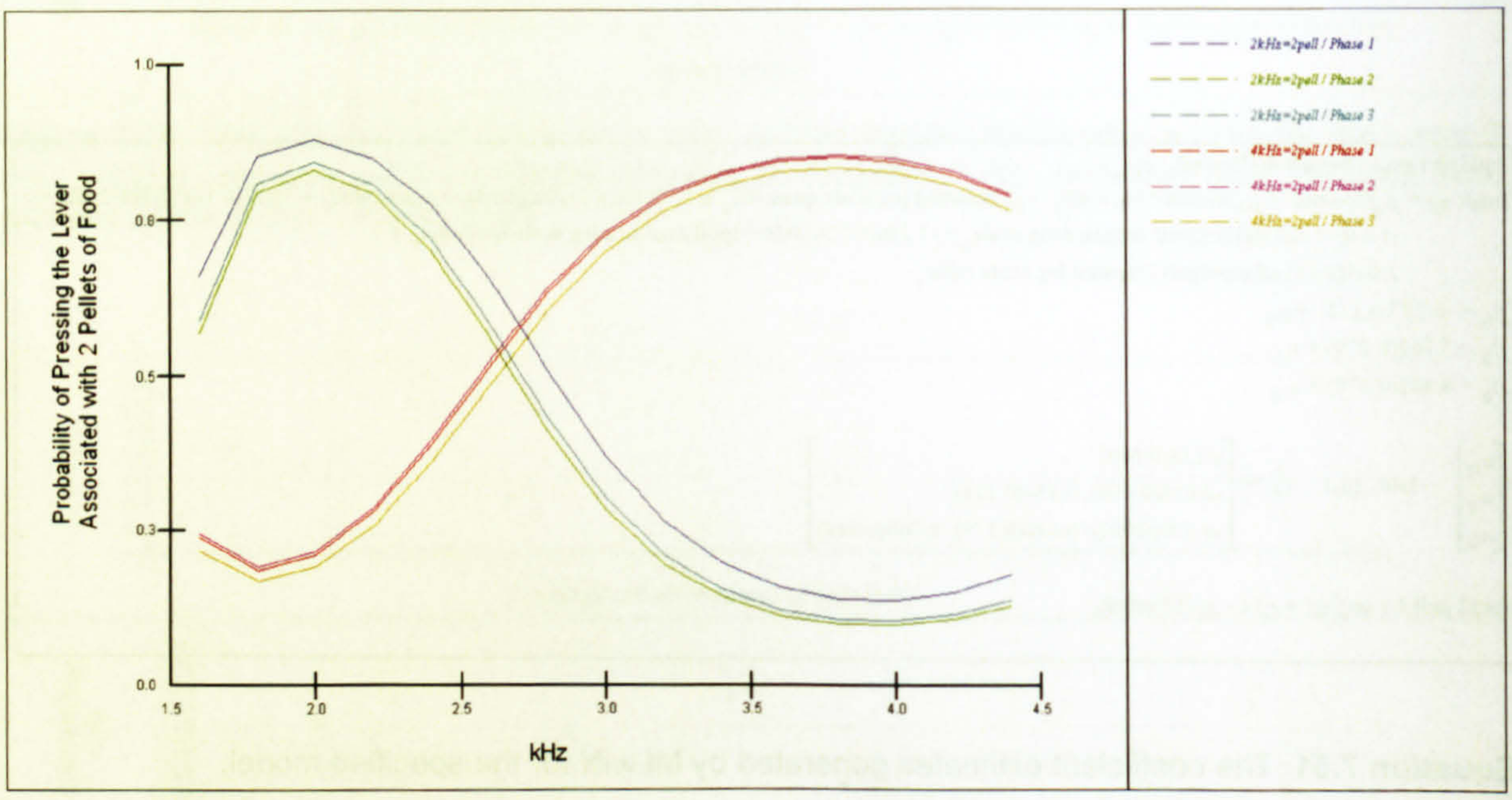

Figure 7.31 The probability of pressing the lever associated with 2 pellets of food (as opposed to the lever associated with 1 pellet of food) across kHz, by *measurement phase / contingency* group. These predicted lines were generated from the model and estimates in Equation 7.53, and indicate the nature of the significant interaction between *measurement phase* and *contingency:* compared to the first *measurement phase,* both *contingency* groups have a lower probability of pressing the '2-pellet' lever in the third *measurement phase*, but only the *2kHz=2pell contingency* group also have a lower probability in the second *measurment phase*.

2 pell lvr $on_{ij} \sim$ Binomial(Constant$_{ij}$, π_{ij})

logit(π_{ij}) = β_{0j}Constant + β_{1j}Standard log scale$_{ij}$ + β_{2j}Standard log scale quadratic$_{ij}$ + -7.831(0.462)Standard log scale cubic$_{ij}$ + 0.656(0.183)4kHz=2pell$_j$ +
-1.464(0.390)4kHz=2pell.Standard log scale$_{ij}$ + -2.167(0.527)4kHz=2pell.Standard log scale quadratic$_{ij}$ +
2.508(0.632)4kHz=2pell.Standard log scale cubic$_{ij}$ + -0.394(0.077)Phase_2$_{ij}$ + -0.304(0.077)Phase_3$_{ij}$ + 0.423(0.106)Phase_2.4kHz=2pell$_{ij}$ +
0.180(0.105)Phase_3.4kHz=2pell$_{ij}$ + 0.274(0.161)Enriched$_j$

β_{0j} = -0.192(0.153) + u_{0j}

β_{1j} = 5.798(0.281) + u_{1j}

β_{2j} = 0.444(0.377) + u_{2j}

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) \quad \Omega_u = \begin{bmatrix} 0.104(0.042) & & \\ -0.015(0.055) & 0.359(0.145) & \\ -0.073(0.090) & 0.348(0.181) & 0.771(0.350) \end{bmatrix}$$

var(2 pell lvr $on_{ij}|\pi_{ij}$) = $\pi_{ij}(1 - \pi_{ij})$/Constant$_{ij}$

Equation 7.54 The coefficient estimates generated by MLwiN for the specified model.

2 pell lvr $on_{ij} \sim$ Binomial(Constant$_{ij}$, π_{ij})

logit(π_{ij}) = β_{0j}Constant + β_{1j}Standard log scale$_{ij}$ + β_{2j}Standard log scale quadratic$_{ij}$ + -7.836(0.462)Standard log scale cubic$_{ij}$ + 0.656(0.183)4kHz=2pell$_j$ +
-1.465(0.390)4kHz=2pell.Standard log scale$_{ij}$ + -2.167(0.527)4kHz=2pell.Standard log scale quadratic$_{ij}$ +
2.509(0.632)4kHz=2pell.Standard log scale cubic$_{ij}$ + -0.253(0.093)Phase_2$_{ij}$ + -0.233(0.093)Phase_3$_{ij}$ + 0.422(0.106)4kHz=2pell.Phase_2$_{ij}$ +
0.180(0.105)4kHz=2pell.Phase_3$_{ij}$ + 0.415(0.172)Enriched$_j$ + -0.284(0.106)Enriched.Phase_2$_{ij}$ + -0.144(0.105)Enriched.Phase_3$_{ij}$

β_{0j} = -0.261(0.156) + u_{0j}

β_{1j} = 5.801(0.281) + u_{1j}

β_{2j} = 0.443(0.377) + u_{2j}

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) \quad \Omega_u = \begin{bmatrix} 0.104(0.042) & & \\ -0.015(0.055) & 0.361(0.145) & \\ -0.073(0.090) & 0.348(0.181) & 0.773(0.350) \end{bmatrix}$$

var(2 pell lvr $on_{ij}|\pi_{ij}$) = $\pi_{ij}(1 - \pi_{ij})$/Constant$_{ij}$

Equation 7.55 The coefficient estimates generated by MLwiN for the specified model.

Ordered Subgroup	Observed (O)	Expected (E) (i.e. predicted from model)	$(O-E)^2$	$(O-E)^2$ / E
1	139	122.82	261.91	2.13
2	183	190.83	61.30	0.32
3	258	260.53	6.40	0.02
4	378	373.21	22.95	0.06
5	531	564.34	1111.68	1.97
6	780	775.94	16.50	0.02
7	949	912.58	1326.45	1.45
8	980	1009.22	853.93	0.85
9	1060	1073.21	174.53	0.16
10	1178	1156.49	462.59	0.40
			X^2 statistic:	**7.39**
			d.f.	**10**
			p	**0.688**

Table 7.5 Calculations pertaining to the Hosmer-Lemeshow goodness-of-fit test of the model specified in Equation 7.55; the 'Observed' and 'Expected' refer to the number of presses on the lever associated with 2 pellets of food, and the subgroups are ordered with respect to the 'Expected' values (with subgroup 1 having the lowest expected values, and subgroup 10 the greatest, with a roughly equal number of *trials* in each subgroup).

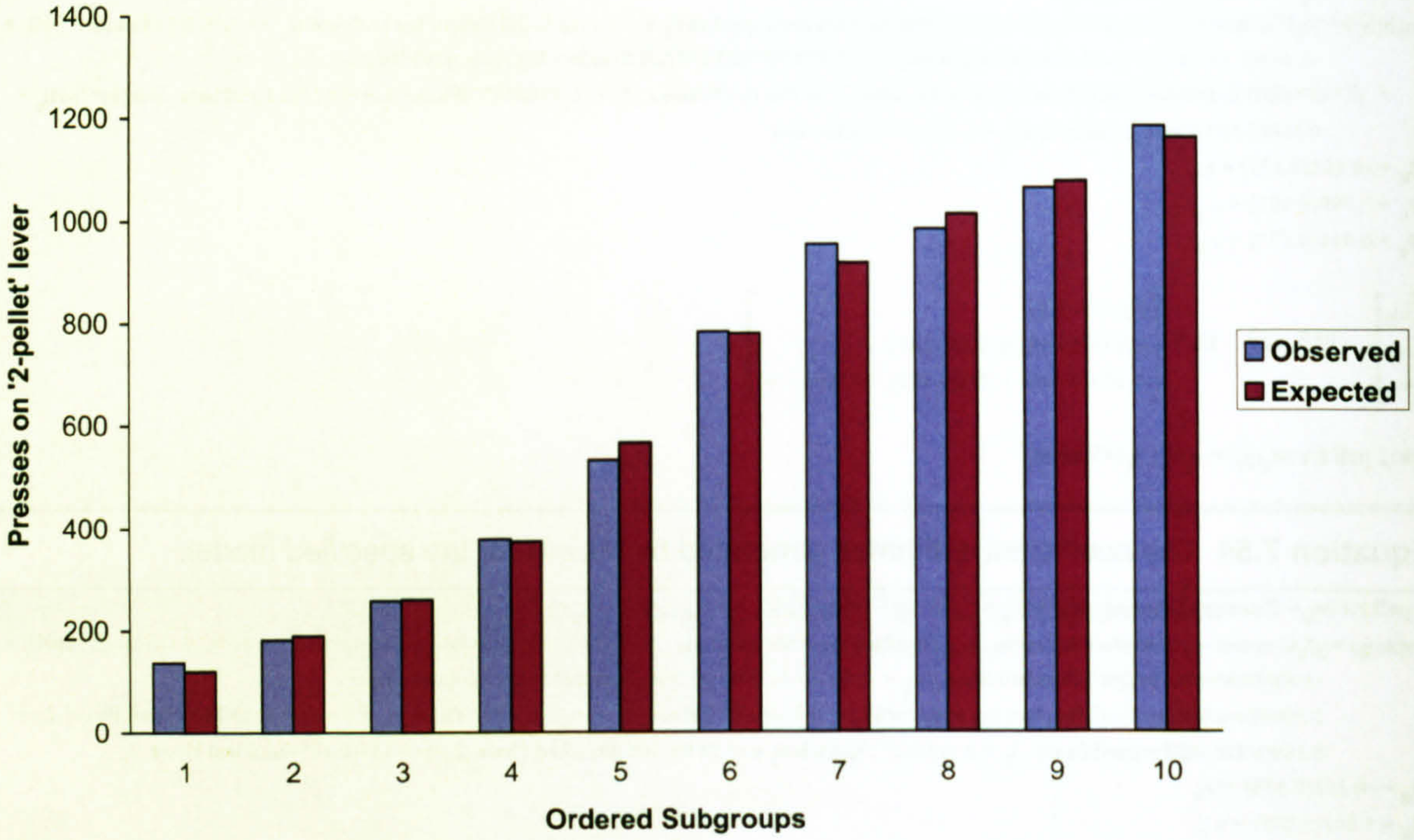

Figure 7.32 Plot of the 'Observed' and 'Expected' number of presses on the lever associated with 2 pellets of food, across the ordered subgroups (see Table 7.5).

$$
\begin{aligned}
&\text{2 pell lvr on}_{ij} \sim \text{Binomial}(\text{Constant}_{ij}, \pi_{ij}) \\
&\text{logit}(\pi_{ij}) = \beta_{0j}\text{Constant} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -7.804(0.461)\text{Standard log scale cubic}_{ij} + 0.857(0.172)\text{4kHz=2pell}_{j} + \\
&\qquad -1.438(0.389)\text{4kHz=2pell.Standard log scale}_{ij} + -2.168(0.526)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + \\
&\qquad 2.473(0.632)\text{4kHz=2pell.Standard log scale cubic}_{ij} + -0.030(0.074)\text{Phase\_2}_{ij} + -0.138(0.074)\text{Phase\_3}_{ij} + 0.416(0.172)\text{Enriched}_{j} + \\
&\qquad -0.285(0.106)\text{Enriched.Phase\_2}_{ij} + -0.145(0.106)\text{Enriched.Phase\_3}_{ij} \\
&\beta_{0j} = -0.367(0.152) + u_{0j} \\
&\beta_{1j} = 5.778(0.280) + u_{1j} \\
&\beta_{2j} = 0.442(0.376) + u_{2j} \\
&\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) \; : \; \Omega_u = \begin{bmatrix} 0.104(0.042) & & \\ -0.015(0.055) & 0.358(0.145) & \\ -0.073(0.090) & 0.347(0.180) & 0.769(0.349) \end{bmatrix} \\
&\text{var}(\text{2 pell lvr on}_{ij} | \pi_{ij}) = \pi_{ij}(1 - \pi_{ij})/\text{Constant}_{ij}
\end{aligned}
$$

Equation 7.56 The coefficient estimates generated by MLwiN for the specified model.

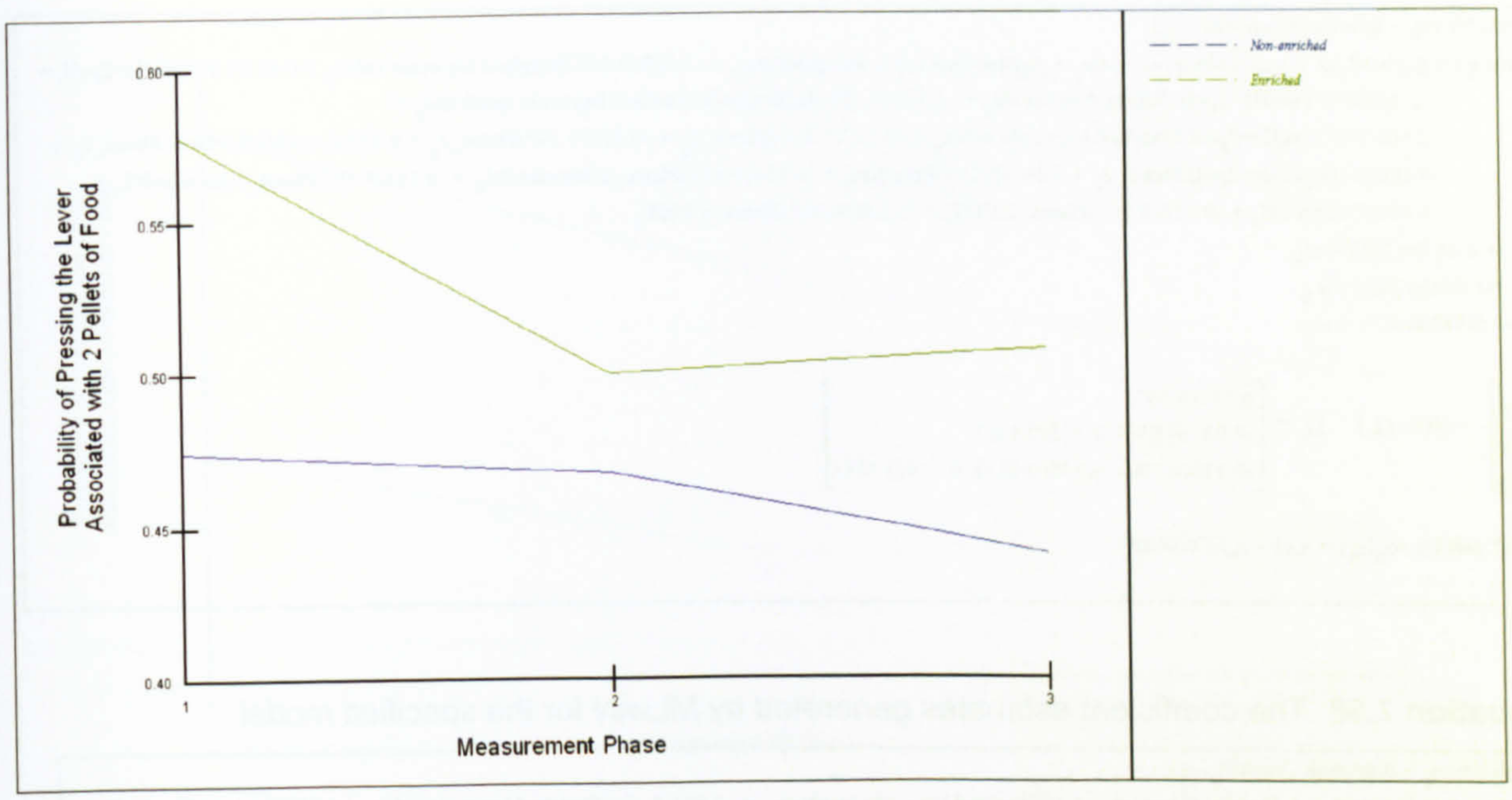

Figure 7.33 The predicted probability of pressing the '2-pellet' lever in the single-frequency probe sessions, by *treatment* group, across *measurement phase.* These predictions are derived from the model specified in Equation 7.56.[183]

$$
\begin{aligned}
&\text{2 pell lvr on}_{ij} \sim \text{Binomial}(\text{Constant}_{ij}, \pi_{ij}) \\
&\text{logit}(\pi_{ij}) = \beta_{0j}\text{Constant} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + \text{-7.839(0.462)Standard log scale cubic}_{ij} + \text{0.657(0.174)4kHz=2pell}_{j} + \\
&\qquad \text{-1.460(0.390)4kHz=2pell.Standard log scale}_{ij} + \text{-2.177(0.528)4kHz=2pell.Standard log scale quadratic}_{ij} + \\
&\qquad \text{2.501(0.632)4kHz=2pell.Standard log scale cubic}_{ij} + \text{-0.253(0.093)Phase\_2}_{ij} + \text{-0.233(0.093)Phase\_3}_{ij} + \text{0.422(0.106)Phase\_2.4kHz=2pell}_{ij} + \\
&\qquad \text{0.179(0.106)Phase\_3.4kHz=2pell}_{ij} + \text{0.410(0.154)Enriched}_{j} + \text{-0.284(0.106)Enriched.Phase\_2}_{ij} + \text{-0.144(0.105)Enriched.Phase\_3}_{ij} + \\
&\qquad \text{0.324(0.141)UHT}_{j} \\
&\beta_{0j} = \text{-0.421(0.162)} + u_{0j} \\
&\beta_{1j} = \text{5.801(0.281)} + u_{1j} \\
&\beta_{2j} = \text{0.446(0.377)} + u_{2j} \\
&\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} \text{0.091(0.037)} & & \\ \text{-0.037(0.053)} & \text{0.360(0.145)} & \\ \text{-0.118(0.090)} & \text{0.347(0.181)} & \text{0.776(0.352)} \end{bmatrix} \\
&\text{var}(\text{2 pell lvr on}_{ij} | \pi_{ij}) = \pi_{ij}(1 - \pi_{ij})/\text{Constant}_{ij}
\end{aligned}
$$

Equation 7.57 The coefficient estimates generated by MLwiN for the specified model.

[183] Specifically, predictions are generated from an equation which includes only those terms represented in the chart: namely *treatment, measurement phase,* and the *treatment*measurement phase* interactions. Rather than plot the resulting predictions as deviations from a reference point of 0.5, we attempt to introduce a more accurate representation of actual probability by taking the average predicted value from an equation which includes all the *other* predictor terms, and *add* this to our original predicted values, prior to antilogit transformation (N.B. we always exclude *subject-level* variance when generating these predictions).

2 pell lvr $on_{ij} \sim$ Binomial(Constant$_{ij}$, π_{ij})

logit(π_{ij}) = β_{0j}Constant + β_{1j}Standard log scale$_{ij}$ + β_{2j}Standard log scale quadratic$_{ij}$ + -7.840(0.462)Standard log scale cubic$_{ij}$ + 0.657(0.174)4kHz=2pell$_j$ + -1.461(0.391)4kHz=2pell.Standard log scale$_{ij}$ + -2.179(0.528)4kHz=2pell.Standard log scale quadratic$_{ij}$ + 2.500(0.632)4kHz=2pell.Standard log scale cubic$_{ij}$ + -0.227(0.106)Phase_2$_{ij}$ + -0.122(0.105)Phase_3$_{ij}$ + 0.422(0.106)4kHz=2pell.Phase_2$_{ij}$ + 0.180(0.106)4kHz=2pell.Phase_3$_{ij}$ + 0.409(0.154)Enriched$_j$ + -0.284(0.106)Phase_2.Enriched#1$_{ij}$ + -0.141(0.105)Phase_3.Enriched#1$_{ij}$ + 0.419(0.154)UHT$_j$ + -0.055(0.106)Phase_2.UHT$_{ij}$ + -0.231(0.105)Phase_3.UHT$_{ij}$

β_{0j} = -0.467(0.164) + u_{0j}

β_{1j} = 5.804(0.282) + u_{1j}

β_{2j} = 0.447(0.377) + u_{2j}

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) \; : \; \Omega_u = \begin{bmatrix} 0.091(0.037) & & \\ -0.037(0.053) & 0.362(0.145) & \\ -0.118(0.090) & 0.348(0.181) & 0.774(0.351) \end{bmatrix}$$

var(2 pell lvr $on_{ij}|\pi_{ij}$) = $\pi_{ij}(1 - \pi_{ij})$/Constant$_{ij}$

Equation 7.58 The coefficient estimates generated by MLwiN for the specified model.

2 pell lvr $on_{ij} \sim$ Binomial(Constant$_{ij}$, π_{ij})

logit(π_{ij}) = β_{0j}Constant + β_{1j}Standard log scale$_{ij}$ + β_{2j}Standard log scale quadratic$_{ij}$ + -7.803(0.461)Standard log scale cubic$_{ij}$ + 0.858(0.163)4kHz=2pell$_j$ + -1.432(0.390)4kHz=2pell.Standard log scale$_{ij}$ + -2.178(0.526)4kHz=2pell.Standard log scale quadratic$_{ij}$ + 2.462(0.631)4kHz=2pell.Standard log scale cubic$_{ij}$ + -0.142(0.074)Phase_2$_{ij}$ + -0.096(0.073)Phase_3$_{ij}$ + 0.268(0.141)Enriched$_j$ + 0.422(0.154)UHT$_j$ + -0.060(0.106)UHT.Phase_2$_{ij}$ + -0.235(0.106)UHT.Phase_3$_{ij}$

β_{0j} = -0.504(0.158) + u_{0j}

β_{1j} = 5.776(0.280) + u_{1j}

β_{2j} = 0.446(0.376) + u_{2j}

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) \; : \; \Omega_u = \begin{bmatrix} 0.091(0.037) & & \\ -0.037(0.052) & 0.359(0.144) & \\ -0.118(0.089) & 0.348(0.180) & 0.768(0.349) \end{bmatrix}$$

var(2 pell lvr $on_{ij}|\pi_{ij}$) = $\pi_{ij}(1 - \pi_{ij})$/Constant$_{ij}$

Equation 7.59 The coefficient estimates generated by MLwiN for the specified model.

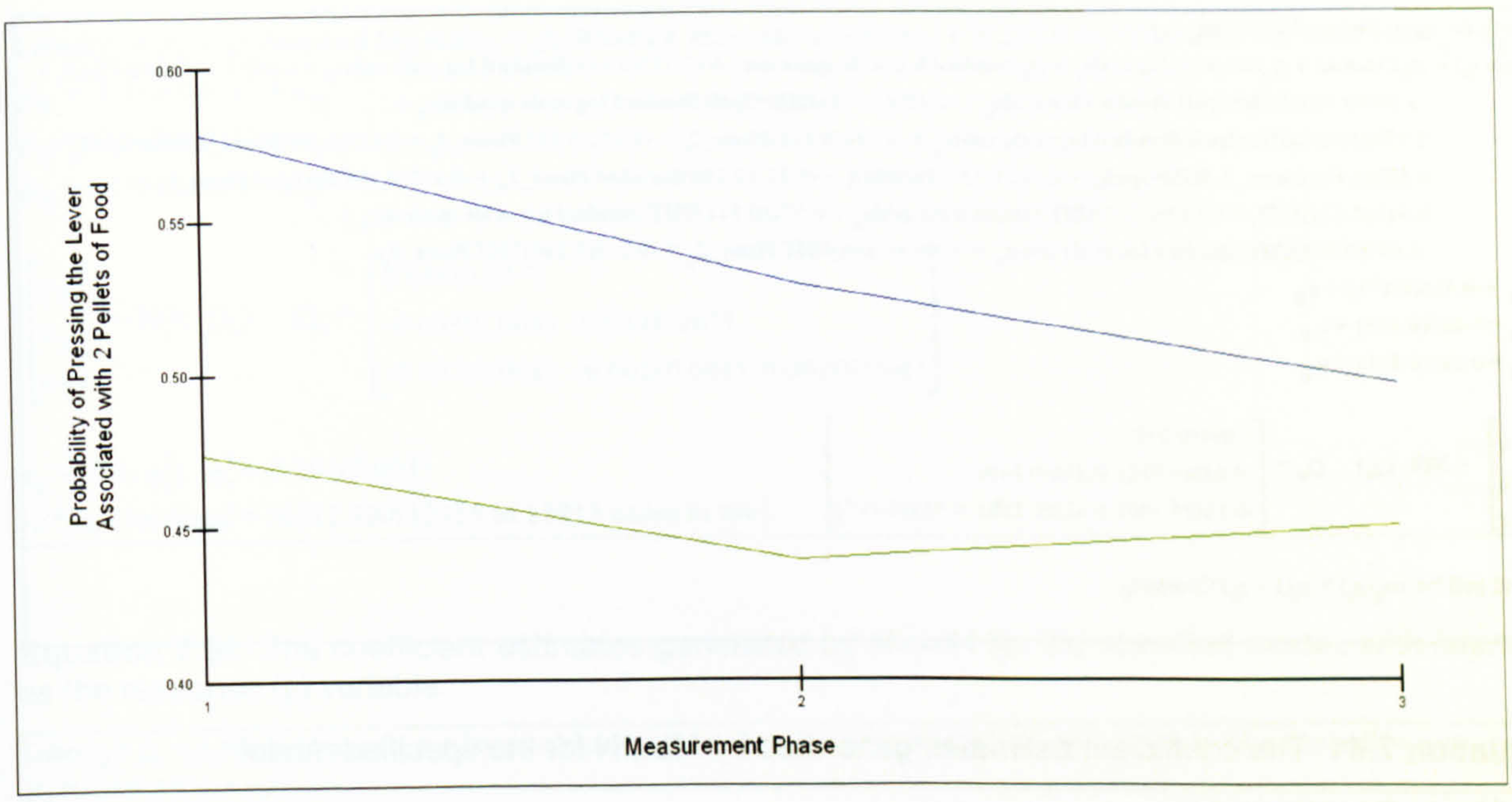

Figure 7.34 The predicted probability of pressing the '2-pellet' lever in the single-frequency probe sessions, by *prior treatment* group, across *measurement phase*. Blue: *UHT* group; Green: *Control* group. These predictions are derived from the model specified in Equation 7.59.

$$
\begin{aligned}
&\text{2 pell lvr on}_{ij} \sim \text{Binomial}(\text{Constant}_{ij}, \pi_{ij}) \\
&\text{logit}(\pi_{ij}) = \beta_{0j}\text{Constant} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + \text{-7.041(0.543)Standard log scale cubic}_{ij} + \text{0.658(0.176)4kHz=2pell}_{j} + \\
&\quad \text{-1.499(0.385)4kHz=2pell.Standard log scale}_{ij} + \text{-2.177(0.570)4kHz=2pell.Standard log scale quadratic}_{ij} + \\
&\quad \text{2.580(0.636)4kHz=2pell.Standard log scale cubic}_{ij} + \text{-0.255(0.093)Phase\_2}_{ij} + \text{-0.234(0.093)Phase\_3}_{ij} + \text{0.425(0.106)Phase\_2.4kHz=2pell}_{ij} + \\
&\quad \text{0.180(0.106)Phase\_3.4kHz=2pell}_{ij} + \text{0.413(0.153)Enriched}_{j} + \text{-0.285(0.106)Enriched.Phase\_2}_{ij} + \text{-0.144(0.106)Enriched.Phase\_3}_{ij} + \\
&\quad \text{0.244(0.164)UHT}_{j} + \text{0.871(0.381)UHT.Standard log scale}_{ij} + \text{0.578(0.543)UHT.Standard log scale quadratic}_{ij} + \\
&\quad \text{-1.673(0.598)UHT.Standard log scale cubic}_{ij} \\
&\beta_{0j} = \text{-0.378(0.168)} + u_{0j} \\
&\beta_{1j} = \text{5.409(0.333)} + u_{1j} \\
&\beta_{2j} = \text{0.143(0.490)} + u_{2j} \\
&\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} \text{0.094(0.038)} & & \\ \text{-0.032(0.052)} & \text{0.340(0.140)} & \\ \text{-0.145(0.100)} & \text{0.328(0.188)} & \text{0.958(0.417)} \end{bmatrix} \\
&\text{var(2 pell lvr on}_{ij}|\pi_{ij}) = \pi_{ij}(1 - \pi_{ij})/\text{Constant}_{ij}
\end{aligned}
$$

Equation 7.60 The coefficient estimates generated by MLwiN for the specified model.

2 pell lvr $on_{ij} \sim$ Binomial(Constant$_{ij}$, π_{ij})

logit(π_{ij}) = β_{0j}Constant + β_{1j}Standard log scale$_{ij}$ + β_{2j}Standard log scale quadratic$_{ij}$ + -7.032(0.543)Standard log scale cubic$_{ij}$ + 0.658(0.176)4kHz=2pell$_j$ + -1.500(0.386)4kHz=2pell.Standard log scale$_{ij}$ + -2.180(0.571)4kHz=2pell.Standard log scale quadratic$_{ij}$ + 2.579(0.636)4kHz=2pell.Standard log scale cubic$_{ij}$ + -0.226(0.106)Phase_2$_{ij}$ + -0.122(0.105)Phase_3$_{ij}$ + 0.425(0.106)Phase_2.4kHz=2pell$_{ij}$ + 0.182(0.106)Phase_3.4kHz=2pell$_{ij}$ + 0.412(0.153)Enriched$_j$ + -0.285(0.106)Enriched.Phase_2$_{ij}$ + -0.142(0.106)Enriched.Phase_3$_{ij}$ + 0.345(0.176)UHT$_j$ + 0.889(0.382)UHT.Standard log scale$_{ij}$ + 0.574(0.544)UHT.Standard log scale quadratic$_{ij}$ + -1.697(0.598)UHT.Standard log scale cubic$_{ij}$ + -0.060(0.106)UHT.Phase_2$_{ij}$ + -0.238(0.106)UHT.Phase_3$_{ij}$

β_{0j} = -0.426(0.171) + u_{0j}

β_{1j} = 5.403(0.333) + u_{1j}

β_{2j} = 0.146(0.491) + u_{2j}

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) \; : \; \Omega_u = \begin{bmatrix} 0.094(0.038) & & \\ -0.032(0.052) & 0.342(0.140) & \\ -0.146(0.100) & 0.329(0.188) & 0.959(0.417) \end{bmatrix}$$

var(2 pell lvr $on_{ij}|\pi_{ij}$) = $\pi_{ij}(1 - \pi_{ij})$/Constant$_{ij}$

Equation 7.61 The coefficient estimates generated by MLwiN for the specified model.

2 pell lvr $on_{ij} \sim$ Binomial(Constant$_{ij}$, π_{ij})

logit(π_{ij}) = β_{0j}Constant + β_{1j}Standard log scale$_{ij}$ + β_{2j}Standard log scale quadratic$_{ij}$ + -7.037(0.544)Standard log scale cubic$_{ij}$ + 0.659(0.174)4kHz=2pell$_j$ + -1.499(0.387)4kHz=2pell.Standard log scale$_{ij}$ + -2.184(0.571)4kHz=2pell.Standard log scale quadratic$_{ij}$ + 2.582(0.637)4kHz=2pell.Standard log scale cubic$_{ij}$ + -0.226(0.106)Phase_2$_{ij}$ + -0.122(0.105)Phase_3$_{ij}$ + 0.425(0.106)Phase_2.4kHz=2pell$_{ij}$ + 0.182(0.106)Phase_3.4kHz=2pell$_{ij}$ + 0.413(0.150)Enriched$_j$ + -0.285(0.106)Enriched.Phase_2$_{ij}$ + -0.142(0.106)Enriched.Phase_3$_{ij}$ + 0.345(0.173)UHT$_j$ + 0.889(0.382)UHT.Standard log scale$_{ij}$ + 0.575(0.544)UHT.Standard log scale quadratic$_{ij}$ + -1.699(0.598)UHT.Standard log scale cubic$_{ij}$ + -0.060(0.106)UHT.Phase_2$_{ij}$ + -0.238(0.106)UHT.Phase_3$_{ij}$ + 0.118(0.137)R$_j$

β_{0j} = -0.485(0.181) + u_{0j}

β_{1j} = 5.407(0.334) + u_{1j}

β_{2j} = 0.150(0.491) + u_{2j}

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) \; : \; \Omega_u = \begin{bmatrix} 0.090(0.037) & & \\ -0.022(0.051) & 0.343(0.141) & \\ -0.140(0.099) & 0.330(0.189) & 0.962(0.418) \end{bmatrix}$$

var(2 pell lvr $on_{ij}|\pi_{ij}$) = $\pi_{ij}(1 - \pi_{ij})$/Constant$_{ij}$

Equation 7.62 The coefficient estimates generated by MLwiN for the specified model.

2 pell lvr $on_{ij} \sim$ Binomial(Constant$_{ij}$, π_{ij})

logit(π_{ij}) = β_{0j}Constant + β_{1j}Standard log scale$_{ij}$ + β_{2j}Standard log scale quadratic$_{ij}$ + -7.055(0.543)Standard log scale cubic$_{ij}$ + 0.639(0.175)4kHz=2pell$_j$ + -1.511(0.385)4kHz=2pell.Standard log scale$_{ij}$ + -2.211(0.564)4kHz=2pell.Standard log scale quadratic$_{ij}$ + 2.563(0.636)4kHz=2pell.Standard log scale cubic$_{ij}$ + -0.242(0.106)Phase_2$_{ij}$ + -0.143(0.105)Phase_3$_{ij}$ + 0.434(0.107)4kHz=2pell.Phase_2$_{ij}$ + 0.204(0.106)4kHz=2pell.Phase_3$_{ij}$ + 0.412(0.149)Enriched$_j$ + -0.290(0.106)Phase_2.Enriched#2$_{ij}$ + -0.146(0.106)Phase_3.Enriched#2$_{ij}$ + 0.328(0.174)UHT$_j$ + 0.876(0.380)UHT.Standard log scale$_{ij}$ + 0.554(0.536)UHT.Standard log scale quadratic$_{ij}$ + -1.655(0.597)UHT.Standard log scale cubic$_{ij}$ + -0.053(0.106)UHT.Phase_2$_{ij}$ + -0.233(0.106)UHT.Phase_3$_{ij}$ + -0.020(0.005)Latency$_{ij}$

β_{0j} = -0.408(0.168) + u_{0j}

β_{1j} = 5.415(0.332) + u_{1j}

β_{2j} = 0.199(0.484) + u_{2j}

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) \; : \; \Omega_u = \begin{bmatrix} 0.092(0.038) & & \\ -0.032(0.052) & 0.338(0.139) & \\ -0.151(0.099) & 0.312(0.184) & 0.926(0.406) \end{bmatrix}$$

var(2 pell lvr $on_{ij}|\pi_{ij}$) = $\pi_{ij}(1 - \pi_{ij})$/Constant$_{ij}$

Equation 7.63 The coefficient estimates generated by MLwiN for the specified model.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + 0.282(0.049)\text{Standard log scale cubic}_{ij} + e_{ij}$$

$$\beta_{0j} = -1.381(0.054) + u_{0j}$$

$$\beta_{1j} = -0.212(0.041) + u_{1j}$$

$$\beta_{2j} = 0.298(0.025) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.046(0.017) & & \\ -0.003(0.008) & 0.020(0.007) & \\ -0.007(0.006) & -0.002(0.004) & 0.003(0.004) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.203(0.003)$$

$$-2 * \textit{loglikelihood} = 16211.390(12915 \text{ of } 12915 \text{ cases in use})$$

Equation 7.64 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (y) variable.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + 0.215(0.050)\text{Standard log scale cubic}_{ij} + -0.058(0.010)\text{2pell lvr press}_{ij} + e_{ij}$$

$$\beta_{0j} = -1.350(0.054) + u_{0j}$$

$$\beta_{1j} = -0.155(0.042) + u_{1j}$$

$$\beta_{2j} = 0.287(0.025) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.046(0.016) & & \\ -0.004(0.008) & 0.020(0.008) & \\ -0.007(0.006) & -0.001(0.004) & 0.002(0.003) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.203(0.003)$$

$$-2 * \textit{loglikelihood} = 16174.850(12915 \text{ of } 12915 \text{ cases in use})$$

Equation 7.65 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (y) variable.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -0.373(0.066)\text{Standard log scale cubic}_{ij} + -0.063(0.012)\text{2pell lvr press}_{ij} + -1.003(0.043)\text{2pell lvr press.Standard log scale}_{ij} + 0.083(0.045)\text{2pell lvr press.Standard log scale quadratic}_{ij} + 1.008(0.098)\text{2pell lvr press.Standard log scale cubic}_{ij} + e_{ij}$$

$$\beta_{0j} = -1.300(0.053) + u_{0j}$$

$$\beta_{1j} = 0.355(0.043) + u_{1j}$$

$$\beta_{2j} = 0.392(0.031) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.044(0.016) & & \\ -0.008(0.007) & 0.015(0.006) & \\ -0.008(0.005) & -0.001(0.003) & 0.000(0.003) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.189(0.002)$$

$$-2 * \textit{loglikelihood} = 15271.270(12915 \text{ of } 12915 \text{ cases in use})$$

Equation 7.66 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (y) variable.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -0.391(0.068)\text{Standard log scale cubic}_{ij} + -0.064(0.012)\text{2pell lvr press}_{ij} +$$
$$-1.003(0.043)\text{2pell lvr press.Standard log scale}_{ij} + 0.087(0.045)\text{2pell lvr press.Standard log scale quadratic}_{ij} +$$
$$1.005(0.098)\text{2pell lvr press.Standard log scale cubic}_{ij} + -0.122(0.086)\text{4kHz=2pell}_{j} + e_{ij}$$

$$\beta_{0j} = -1.238(0.071) + u_{0j}$$
$$\beta_{1j} = 0.361(0.043) + u_{1j}$$
$$\beta_{2j} = 0.389(0.031) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.048(0.017) & & \\ -0.011(0.007) & 0.015(0.006) & \\ -0.009(0.005) & -0.001(0.003) & 0.001(0.003) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.189(0.002)$$

*-2*loglikelihood* = 15269.700(12915 of 12915 cases in use)

Equation 7.67 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -0.609(0.116)\text{Standard log scale cubic}_{ij} + 0.007(0.018)\text{2pell lvr press}_{ij} +$$
$$-1.485(0.068)\text{2pell lvr press.Standard log scale}_{ij} + 0.007(0.084)\text{2pell lvr press.Standard log scale quadratic}_{ij} +$$
$$1.573(0.173)\text{2pell lvr press.Standard log scale cubic}_{ij} + 0.068(0.102)\text{4kHz=2pell}_{j} + -0.357(0.087)\text{4kHz=2pell.Standard log scale}_{ij} +$$
$$-0.443(0.079)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + 0.010(0.168)\text{4kHz=2pell.Standard log scale cubic}_{ij} +$$
$$-0.187(0.026)\text{2pell lvr press.4kHz=2pell}_{ij} + 0.869(0.094)\text{2pell lvr press.4kHz=2pell.Standard log scale}_{ij} +$$
$$0.412(0.117)\text{2pell lvr press.4kHz=2pell.Standard log scale quadratic}_{ij} + -0.839(0.242)\text{2pell lvr press.4kHz=2pell.Standard log scale cubic}_{ij} + e_{ij}$$

$$\beta_{0j} = -1.306(0.072) + u_{0j}$$
$$\beta_{1j} = 0.576(0.063) + u_{1j}$$
$$\beta_{2j} = 0.487(0.044) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.040(0.014) & & \\ -0.008(0.006) & 0.012(0.005) & \\ 0.000(0.000) & 0.000(0.000) & 0.000(0.000) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.186(0.002)$$

*-2*loglikelihood* = 15035.500(12915 of 12915 cases in use)

Equation 7.68 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -0.623(0.116)\text{Standard log scale cubic}_{ij} + 0.003(0.018)\text{2pell lvr press}_{ij} +$$
$$-1.495(0.068)\text{2pell lvr press.Standard log scale}_{ij} + 0.010(0.084)\text{2pell lvr press.Standard log scale quadratic}_{ij} +$$
$$1.584(0.172)\text{2pell lvr press.Standard log scale cubic}_{ij} + 0.064(0.102)\text{4kHz=2pell}_{j} + -0.365(0.086)\text{4kHz=2pell.Standard log scale}_{ij} +$$
$$-0.436(0.079)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + 0.030(0.167)\text{4kHz=2pell.Standard log scale cubic}_{ij} +$$
$$-0.179(0.026)\text{2pell lvr press.4kHz=2pell}_{ij} + 0.874(0.094)\text{2pell lvr press.4kHz=2pell.Standard log scale}_{ij} +$$
$$0.400(0.116)\text{2pell lvr press.4kHz=2pell.Standard log scale quadratic}_{ij} + -0.854(0.241)\text{2pell lvr press.4kHz=2pell.Standard log scale cubic}_{ij} +$$
$$-0.070(0.009)\text{Phase\_2}_{ij} + 0.023(0.009)\text{Phase\_3}_{ij} + e_{ij}$$

$$\beta_{0j} = -1.289(0.072) + u_{0j}$$
$$\beta_{1j} = 0.585(0.063) + u_{1j}$$
$$\beta_{2j} = 0.488(0.043) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.040(0.014) & & \\ -0.008(0.006) & 0.012(0.005) & \\ 0.000(0.000) & 0.000(0.000) & 0.000(0.000) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.184(0.002)$$

*-2*loglikelihood* = 14926.610(12915 of 12915 cases in use)

Equation 7.69 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -0.619(0.115)\text{Standard log scale cubic}_{ij} + 0.003(0.018)\text{2pell lvr press}_{ij} + -1.490(0.068)\text{2pell lvr press.Standard log scale}_{ij} + 0.008(0.084)\text{2pell lvr press.Standard log scale quadratic}_{ij} + 1.580(0.172)\text{2pell lvr press.Standard log scale cubic}_{ij} + 0.052(0.102)\text{4kHz=2pell}_{j} + -0.364(0.086)\text{4kHz=2pell.Standard log scale}_{ij} + -0.433(0.079)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + 0.031(0.167)\text{4kHz=2pell.Standard log scale cubic}_{ij} + -0.178(0.026)\text{2pell lvr press.4kHz=2pell}_{ij} + 0.869(0.094)\text{2pell lvr press.4kHz=2pell.Standard log scale}_{ij} + 0.399(0.116)\text{2pell lvr press.4kHz=2pell.Standard log scale quadratic}_{ij} + -0.853(0.241)\text{2pell lvr press.4kHz=2pell.Standard log scale cubic}_{ij} + -0.061(0.013)\text{Phase\_2}_{ij} + -0.002(0.013)\text{Phase\_3}_{ij} + -0.017(0.019)\text{4kHz=2pell.Phase\_2}_{ij} + 0.051(0.019)\text{4kHz=2pell.Phase\_3}_{ij} + e_{ij}$$

$$\beta_{0j} = -1.283(0.072) + u_{0j}$$
$$\beta_{1j} = 0.583(0.063) + u_{1j}$$
$$\beta_{2j} = 0.488(0.043) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.040(0.014) & & \\ -0.008(0.006) & 0.012(0.005) & \\ 0.000(0.000) & 0.000(0.000) & 0.000(0.000) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.184(0.002)$$

*-2*loglikelihood* = 14912.250(12915 of 12915 cases in use)

Equation 7.70 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -0.619(0.115)\text{Standard log scale cubic}_{ij} + 0.003(0.018)\text{2pell lvr press}_{ij} + -1.490(0.068)\text{2pell lvr press.Standard log scale}_{ij} + 0.008(0.084)\text{2pell lvr press.Standard log scale quadratic}_{ij} + 1.579(0.172)\text{2pell lvr press.Standard log scale cubic}_{ij} + 0.052(0.098)\text{4kHz=2pell}_{j} + -0.364(0.086)\text{4kHz=2pell.Standard log scale}_{ij} + -0.433(0.079)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + 0.031(0.167)\text{4kHz=2pell.Standard log scale cubic}_{ij} + -0.178(0.026)\text{2pell lvr press.4kHz=2pell}_{ij} + 0.869(0.094)\text{2pell lvr press.4kHz=2pell.Standard log scale}_{ij} + 0.399(0.116)\text{2pell lvr press.4kHz=2pell.Standard log scale quadratic}_{ij} + -0.852(0.241)\text{2pell lvr press.4kHz=2pell.Standard log scale cubic}_{ij} + -0.061(0.013)\text{Phase\_2}_{ij} + -0.002(0.013)\text{Phase\_3}_{ij} + -0.017(0.019)\text{4kHz=2pell.Phase\_2}_{ij} + 0.051(0.019)\text{4kHz=2pell.Phase\_3}_{ij} + -0.105(0.090)\text{Enriched}_{j} + e_{ij}$$

$$\beta_{0j} = -1.231(0.082) + u_{0j}$$
$$\beta_{1j} = 0.583(0.063) + u_{1j}$$
$$\beta_{2j} = 0.488(0.043) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.036(0.013) & & \\ -0.007(0.006) & 0.012(0.005) & \\ 0.000(0.000) & 0.000(0.000) & 0.000(0.000) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.184(0.002)$$

*-2*loglikelihood* = 14910.990(12915 of 12915 cases in use)

Equation 7.71 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -0.610(0.115)\text{Standard log scale cubic}_{ij} + -0.051(0.020)\text{2pell lvr press}_{ij} +$$
$$-1.480(0.068)\text{2pell lvr press.Standard log scale}_{ij} + 0.027(0.083)\text{2pell lvr press.Standard log scale quadratic}_{ij} +$$
$$1.545(0.172)\text{2pell lvr press.Standard log scale cubic}_{ij} + 0.047(0.106)\text{4kHz=2pell}_{j} + -0.370(0.086)\text{4kHz=2pell.Standard log scale}_{ij} +$$
$$-0.420(0.080)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + 0.041(0.167)\text{4kHz=2pell.Standard log scale cubic}_{ij} +$$
$$-0.171(0.026)\text{2pell lvr press.4kHz=2pell}_{ij} + 0.872(0.093)\text{2pell lvr press.4kHz=2pell.Standard log scale}_{ij} +$$
$$0.361(0.116)\text{2pell lvr press.4kHz=2pell.Standard log scale quadratic}_{ij} + -0.834(0.240)\text{2pell lvr press.4kHz=2pell.Standard log scale cubic}_{ij} +$$
$$-0.060(0.013)\text{Phase\_2}_{ij} + -0.002(0.013)\text{Phase\_3}_{ij} + -0.016(0.019)\text{4kHz=2pell.Phase\_2}_{ij} + 0.053(0.018)\text{4kHz=2pell.Phase\_3}_{ij} +$$
$$-0.044(0.082)\text{Enriched}_{j} + 0.102(0.018)\text{2pell lvr press.Enriched}_{ij} + e_{ij}$$

$$\beta_{0j} = -1.262(0.086) + u_{0j}$$
$$\beta_{1j} = 0.579(0.062) + u_{1j}$$
$$\beta_{2j} = 0.486(0.044) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.043(0.016) & & \\ -0.005(0.006) & 0.012(0.005) & \\ -0.007(0.005) & -0.004(0.003) & 0.001(0.003) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.184(0.002)$$

*-2*loglikelihood* = 14873.540(12915 of 12915 cases in use)

Equation 7.72 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -0.623(0.115)\text{Standard log scale cubic}_{ij} + 0.002(0.018)\text{2pell lvr press}_{ij} +$$
$$-1.491(0.067)\text{2pell lvr press.Standard log scale}_{ij} + 0.007(0.084)\text{2pell lvr press.Standard log scale quadratic}_{ij} +$$
$$1.582(0.172)\text{2pell lvr press.Standard log scale cubic}_{ij} + 0.050(0.097)\text{4kHz=2pell}_{j} + -0.365(0.086)\text{4kHz=2pell.Standard log scale}_{ij} +$$
$$-0.425(0.079)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + 0.044(0.167)\text{4kHz=2pell.Standard log scale cubic}_{ij} +$$
$$-0.174(0.026)\text{2pell lvr press.4kHz=2pell}_{ij} + 0.866(0.093)\text{2pell lvr press.4kHz=2pell.Standard log scale}_{ij} +$$
$$0.392(0.116)\text{2pell lvr press.4kHz=2pell.Standard log scale quadratic}_{ij} + -0.866(0.240)\text{2pell lvr press.4kHz=2pell.Standard log scale cubic}_{ij} +$$
$$-0.045(0.016)\text{Phase\_2}_{ij} + -0.054(0.016)\text{Phase\_3}_{ij} + -0.017(0.018)\text{4kHz=2pell.Phase\_2}_{ij} + 0.051(0.018)\text{4kHz=2pell.Phase\_3}_{ij} +$$
$$-0.128(0.091)\text{Enriched}_{j} + -0.034(0.018)\text{Enriched.Phase\_2}_{ij} + 0.103(0.018)\text{Enriched.Phase\_3}_{ij} + e_{ij}$$

$$\beta_{0j} = -1.219(0.082) + u_{0j}$$
$$\beta_{1j} = 0.585(0.063) + u_{1j}$$
$$\beta_{2j} = 0.489(0.043) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.036(0.013) & & \\ -0.007(0.006) & 0.012(0.005) & \\ 0.000(0.000) & 0.000(0.000) & 0.000(0.000) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.183(0.002)$$

*-2*loglikelihood* = 14851.200(12915 of 12915 cases in use)

Equation 7.73 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -0.627(0.115)\text{Standard log scale cubic}_{ij} + 0.001(0.018)\text{2pell lvr press}_{ij} + -1.495(0.068)\text{2pell lvr press.Standard log scale}_{ij} + 0.009(0.084)\text{2pell lvr press.Standard log scale quadratic}_{ij} + 1.587(0.172)\text{2pell lvr press.Standard log scale cubic}_{ij} + 0.062(0.097)\text{4kHz=2pell}_{j} + -0.367(0.086)\text{4kHz=2pell.Standard log scale}_{ij} + -0.428(0.079)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + 0.043(0.167)\text{4kHz=2pell.Standard log scale cubic}_{ij} + -0.176(0.026)\text{2pell lvr press.4kHz=2pell}_{ij} + 0.871(0.093)\text{2pell lvr press.4kHz=2pell.Standard log scale}_{ij} + 0.394(0.116)\text{2pell lvr press.4kHz=2pell.Standard log scale quadratic}_{ij} + -0.867(0.240)\text{2pell lvr press.4kHz=2pell.Standard log scale cubic}_{ij} + -0.053(0.013)\text{Phase\_2}_{ij} + -0.028(0.013)\text{Phase\_3}_{ij} + -0.128(0.091)\text{Enriched}_{j} + -0.034(0.018)\text{Enriched.Phase\_2}_{ij} + 0.103(0.018)\text{Enriched.Phase\_3}_{ij} + e_{ij}$$

$$\beta_{0j} = -1.224(0.082) + u_{0j}$$

$$\beta_{1j} = 0.587(0.063) + u_{1j}$$

$$\beta_{2j} = 0.489(0.043) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0,\ \Omega_u) \quad \Omega_u = \begin{bmatrix} 0.036(0.013) & & \\ -0.007(0.006) & 0.012(0.005) & \\ 0.000(0.000) & 0.000(0.000) & 0.000(0.000) \end{bmatrix}$$

$$e_{ij} \sim N(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.183(0.002)$$

$-2*\text{loglikelihood} = 14865.610(12915 \text{ of } 12915 \text{ cases in use})$

Equation 7.74 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable.

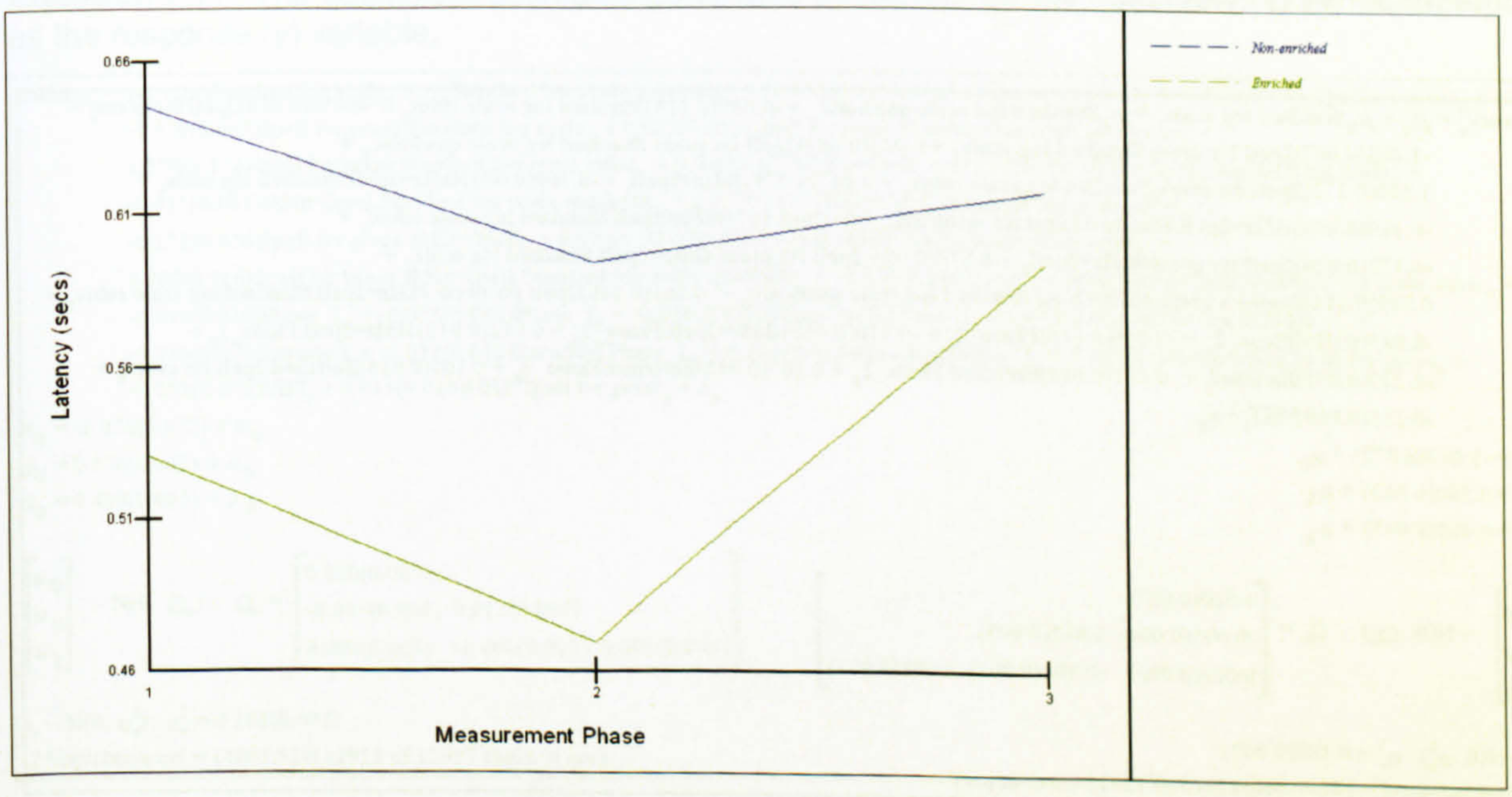

Figure 7.35 The predicted effect of *treatment* on the *latency* to press either lever, across different levels of *measurement phase* (from a prediction equation derived from the model specified in Equation 7.74).

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + \text{-0.613(0.115)Standard log scale cubic}_{ij} + \text{-0.053(0.020)2pell lvr press}_{ij} +$$
$$\text{-1.481(0.067)2pell lvr press.Standard log scale}_{ij} + \text{0.027(0.083)2pell lvr press.Standard log scale quadratic}_{ij} +$$
$$\text{1.547(0.171)2pell lvr press.Standard log scale cubic}_{ij} + \text{0.045(0.107)4kHz=2pell}_{j} + \text{-0.372(0.086)4kHz=2pell.Standard log scale}_{ij} +$$
$$\text{-0.412(0.080)4kHz=2pell.Standard log scale quadratic}_{ij} + \text{0.054(0.166)4kHz=2pell.Standard log scale cubic}_{ij} +$$
$$\text{-0.167(0.026)2pell lvr press.4kHz=2pell}_{ij} + \text{0.869(0.093)2pell lvr press.4kHz=2pell.Standard log scale}_{ij} +$$
$$\text{0.354(0.115)2pell lvr press.4kHz=2pell.Standard log scale quadratic}_{ij} + \text{-0.847(0.239)2pell lvr press.4kHz=2pell.Standard log scale cubic}_{ij} +$$
$$\text{-0.045(0.016)Phase\_2}_{ij} + \text{-0.056(0.016)Phase\_3}_{ij} + \text{-0.016(0.018)4kHz=2pell.Phase\_2}_{ij} + \text{0.053(0.018)4kHz=2pell.Phase\_3}_{ij} +$$
$$\text{-0.070(0.083)Enriched}_{j} + \text{-0.031(0.018)Enriched.Phase\_2}_{ij} + \text{0.107(0.018)Enriched.Phase\_3}_{ij} + \text{0.104(0.018)Enriched.2pell lvr press}_{ij} + e_{ij}$$

$$\beta_{0j} = \text{-1.248(0.086)} + u_{0j}$$
$$\beta_{1j} = \text{0.581(0.062)} + u_{1j}$$
$$\beta_{2j} = \text{0.486(0.045)} + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} \text{0.044(0.016)} & & \\ \text{-0.005(0.006)} & \text{0.012(0.005)} & \\ \text{-0.007(0.005)} & \text{-0.004(0.003)} & \text{0.001(0.003)} \end{bmatrix}$$

$$e_{ij} \sim N(0, \sigma_e^2) \quad \sigma_e^2 = \text{0.183(0.002)}$$

*-2*loglikelihood* = 14812.240(12915 of 12915 cases in use)

Equation 7.75 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + \text{-0.619(0.115)Standard log scale cubic}_{ij} + \text{-0.052(0.020)2pell lvr press}_{ij} +$$
$$\text{-1.482(0.067)2pell lvr press.Standard log scale}_{ij} + \text{0.024(0.083)2pell lvr press.Standard log scale quadratic}_{ij} +$$
$$\text{1.569(0.172)2pell lvr press.Standard log scale cubic}_{ij} + \text{0.047(0.074)4kHz=2pell}_{j} + \text{-0.369(0.086)4kHz=2pell.Standard log scale}_{ij} +$$
$$\text{-0.419(0.080)4kHz=2pell.Standard log scale quadratic}_{ij} + \text{0.050(0.167)4kHz=2pell.Standard log scale cubic}_{ij} +$$
$$\text{-0.172(0.026)2pell lvr press.4kHz=2pell}_{ij} + \text{0.870(0.093)2pell lvr press.4kHz=2pell.Standard log scale}_{ij} +$$
$$\text{0.369(0.116)2pell lvr press.4kHz=2pell.Standard log scale quadratic}_{ij} + \text{-0.861(0.240)2pell lvr press.4kHz=2pell.Standard log scale cubic}_{ij} +$$
$$\text{-0.045(0.016)Phase\_2}_{ij} + \text{-0.056(0.016)Phase\_3}_{ij} + \text{-0.016(0.018)4kHz=2pell.Phase\_2}_{ij} + \text{0.052(0.018)4kHz=2pell.Phase\_3}_{ij} +$$
$$\text{-0.229(0.071)Enriched}_{j} + \text{-0.031(0.018)Enriched.Phase\_2}_{ij} + \text{0.107(0.018)Enriched.Phase\_3}_{ij} + \text{0.104(0.018)Enriched.2pell lvr press}_{ij} +$$
$$\text{-0.253(0.069)UHT}_{j} + e_{ij}$$

$$\beta_{0j} = \text{-1.043(0.072)} + u_{0j}$$
$$\beta_{1j} = \text{0.580(0.062)} + u_{1j}$$
$$\beta_{2j} = \text{0.484(0.045)} + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} \text{0.020(0.007)} & & \\ \text{-0.005(0.004)} & \text{0.012(0.005)} & \\ \text{0.003(0.003)} & \text{-0.004(0.003)} & \text{0.001(0.003)} \end{bmatrix}$$

$$e_{ij} \sim N(0, \sigma_e^2) \quad \sigma_e^2 = \text{0.183(0.002)}$$

*-2*loglikelihood* = 14807.010(12915 of 12915 cases in use)

Equation 7.76 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -0.619(0.115)\text{Standard log scale cubic}_{ij} + -0.051(0.020)\text{2pell lvr press}_{ij} +$$
$$-1.483(0.067)\text{2pell lvr press.Standard log scale}_{ij} + 0.023(0.083)\text{2pell lvr press.Standard log scale quadratic}_{ij} +$$
$$1.571(0.172)\text{2pell lvr press.Standard log scale cubic}_{ij} + 0.048(0.074)\text{4kHz=2pell}_{j} + -0.368(0.086)\text{4kHz=2pell.Standard log scale}_{ij} +$$
$$-0.421(0.080)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + 0.049(0.167)\text{4kHz=2pell.Standard log scale cubic}_{ij} +$$
$$-0.173(0.026)\text{2pell lvr press.4kHz=2pell}_{ij} + 0.870(0.093)\text{2pell lvr press.4kHz=2pell.Standard log scale}_{ij} +$$
$$0.373(0.116)\text{2pell lvr press.4kHz=2pell.Standard log scale quadratic}_{ij} + -0.863(0.240)\text{2pell lvr press.4kHz=2pell.Standard log scale cubic}_{ij} +$$
$$-0.045(0.016)\text{Phase\_2}_{ij} + -0.056(0.016)\text{Phase\_3}_{ij} + -0.017(0.018)\text{4kHz=2pell.Phase\_2}_{ij} + 0.052(0.018)\text{4kHz=2pell.Phase\_3}_{ij} +$$
$$-0.300(0.097)\text{Enriched}_{j} + -0.031(0.018)\text{Enriched.Phase\_2}_{ij} + 0.107(0.018)\text{Enriched.Phase\_3}_{ij} + 0.104(0.018)\text{Enriched.2pell lvr press}_{ij} +$$
$$-0.321(0.096)\text{UHT}_{j} + 0.090(0.136)\text{UHT.Enriched}_{j} + e_{ij}$$

$$\beta_{0j} = -0.996(0.079) + u_{0j}$$
$$\beta_{1j} = 0.580(0.062) + u_{1j}$$
$$\beta_{2j} = 0.484(0.045) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.020(0.007) & & \\ -0.005(0.004) & 0.012(0.005) & \\ 0.004(0.003) & -0.004(0.003) & 0.001(0.003) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.183(0.002)$$

*-2*loglikelihood* = 14806.670(12915 of 12915 cases in use)

Equation 7.77 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -0.616(0.115)\text{Standard log scale cubic}_{ij} + -0.068(0.022)\text{2pell lvr press}_{ij} +$$
$$-1.479(0.067)\text{2pell lvr press.Standard log scale}_{ij} + 0.025(0.083)\text{2pell lvr press.Standard log scale quadratic}_{ij} +$$
$$1.570(0.172)\text{2pell lvr press.Standard log scale cubic}_{ij} + 0.046(0.074)\text{4kHz=2pell}_{j} + -0.370(0.086)\text{4kHz=2pell.Standard log scale}_{ij} +$$
$$-0.411(0.081)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + 0.057(0.167)\text{4kHz=2pell.Standard log scale cubic}_{ij} +$$
$$-0.171(0.026)\text{2pell lvr press.4kHz=2pell}_{ij} + 0.873(0.093)\text{2pell lvr press.4kHz=2pell.Standard log scale}_{ij} +$$
$$0.364(0.116)\text{2pell lvr press.4kHz=2pell.Standard log scale quadratic}_{ij} + -0.871(0.240)\text{2pell lvr press.4kHz=2pell.Standard log scale cubic}_{ij} +$$
$$-0.045(0.016)\text{Phase\_2}_{ij} + -0.055(0.016)\text{Phase\_3}_{ij} + -0.016(0.018)\text{4kHz=2pell.Phase\_2}_{ij} + 0.052(0.018)\text{4kHz=2pell.Phase\_3}_{ij} +$$
$$-0.235(0.070)\text{Enriched}_{j} + -0.031(0.018)\text{Enriched.Phase\_2}_{ij} + 0.106(0.018)\text{Enriched.Phase\_3}_{ij} + 0.105(0.018)\text{Enriched.2pell lvr press}_{ij} +$$
$$-0.282(0.070)\text{UHT}_{j} + 0.033(0.018)\text{UHT.2pell lvr press}_{ij} + e_{ij}$$

$$\beta_{0j} = -1.026(0.072) + u_{0j}$$
$$\beta_{1j} = 0.576(0.062) + u_{1j}$$
$$\beta_{2j} = 0.480(0.045) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.020(0.007) & & \\ -0.005(0.004) & 0.012(0.005) & \\ 0.003(0.003) & -0.004(0.003) & 0.001(0.003) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.183(0.002)$$

*-2*loglikelihood* = 14803.520(12915 of 12915 cases in use)

Equation 7.78 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -0.623(0.115)\text{Standard log scale cubic}_{ij} + -0.052(0.020)\text{2pell lvr press}_{ij} +$$
$$-1.487(0.067)\text{2pell lvr press.Standard log scale}_{ij} + 0.023(0.083)\text{2pell lvr press.Standard log scale quadratic}_{ij} +$$
$$1.577(0.172)\text{2pell lvr press.Standard log scale cubic}_{ij} + 0.047(0.074)\text{4kHz=2pell}_{j} + -0.371(0.086)\text{4kHz=2pell.Standard log scale}_{ij} +$$
$$-0.417(0.080)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + 0.056(0.167)\text{4kHz=2pell.Standard log scale cubic}_{ij} +$$
$$-0.171(0.026)\text{2pell lvr press.4kHz=2pell}_{ij} + 0.874(0.093)\text{2pell lvr press.4kHz=2pell.Standard log scale}_{ij} +$$
$$0.368(0.116)\text{2pell lvr press.4kHz=2pell.Standard log scale quadratic}_{ij} + -0.870(0.240)\text{2pell lvr press.4kHz=2pell.Standard log scale cubic}_{ij} +$$
$$-0.065(0.018)\text{Phase\_2}_{ij} + -0.077(0.018)\text{Phase\_3}_{ij} + -0.016(0.018)\text{4kHz=2pell.Phase\_2}_{ij} + 0.053(0.018)\text{4kHz=2pell.Phase\_3}_{ij} +$$
$$-0.228(0.071)\text{Enriched}_{j} + -0.031(0.018)\text{Enriched.Phase\_2}_{ij} + 0.107(0.018)\text{Enriched.Phase\_3}_{ij} + 0.104(0.018)\text{Enriched.2pell lvr press}_{ij} +$$
$$-0.281(0.069)\text{UHT}_{j} + 0.040(0.018)\text{UHT.Phase\_2}_{ij} + 0.043(0.018)\text{UHT.Phase\_3}_{ij} + e_{ij}$$

$$\beta_{0j} = -1.029(0.072) + u_{0j}$$
$$\beta_{1j} = 0.582(0.062) + u_{1j}$$
$$\beta_{2j} = 0.484(0.045) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.020(0.007) & & \\ -0.005(0.004) & 0.012(0.005) & \\ 0.003(0.003) & -0.004(0.003) & 0.001(0.003) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.183(0.002)$$

*-2*loglikelihood* = 14800.270(12915 of 12915 cases in use)

Equation 7.79 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -0.627(0.115)\text{Standard log scale cubic}_{ij} + 0.003(0.018)\text{2pell lvr press}_{ij} +$$
$$-1.498(0.068)\text{2pell lvr press.Standard log scale}_{ij} + 0.010(0.084)\text{2pell lvr press.Standard log scale quadratic}_{ij} +$$
$$1.592(0.172)\text{2pell lvr press.Standard log scale cubic}_{ij} + 0.064(0.075)\text{4kHz=2pell}_{j} + -0.367(0.086)\text{4kHz=2pell.Standard log scale}_{ij} +$$
$$-0.434(0.079)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + 0.037(0.167)\text{4kHz=2pell.Standard log scale cubic}_{ij} +$$
$$-0.179(0.026)\text{2pell lvr press.4kHz=2pell}_{ij} + 0.877(0.094)\text{2pell lvr press.4kHz=2pell.Standard log scale}_{ij} +$$
$$0.398(0.116)\text{2pell lvr press.4kHz=2pell.Standard log scale quadratic}_{ij} + -0.863(0.241)\text{2pell lvr press.4kHz=2pell.Standard log scale cubic}_{ij} +$$
$$-0.090(0.013)\text{Phase\_2}_{ij} + 0.001(0.013)\text{Phase\_3}_{ij} + -0.109(0.068)\text{Enriched}_{j} + -0.263(0.069)\text{UHT}_{j} + 0.041(0.019)\text{UHT.Phase\_2}_{ij} +$$
$$0.044(0.019)\text{UHT.Phase\_3}_{ij} + e_{ij}$$

$$\beta_{0j} = -1.102(0.072) + u_{0j}$$
$$\beta_{1j} = 0.586(0.063) + u_{1j}$$
$$\beta_{2j} = 0.488(0.043) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.021(0.008) & & \\ -0.006(0.004) & 0.012(0.005) & \\ 0.000(0.000) & 0.000(0.000) & 0.000(0.000) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.184(0.002)$$

*-2*loglikelihood* = 14909.620(12915 of 12915 cases in use)

Equation 7.80 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable.

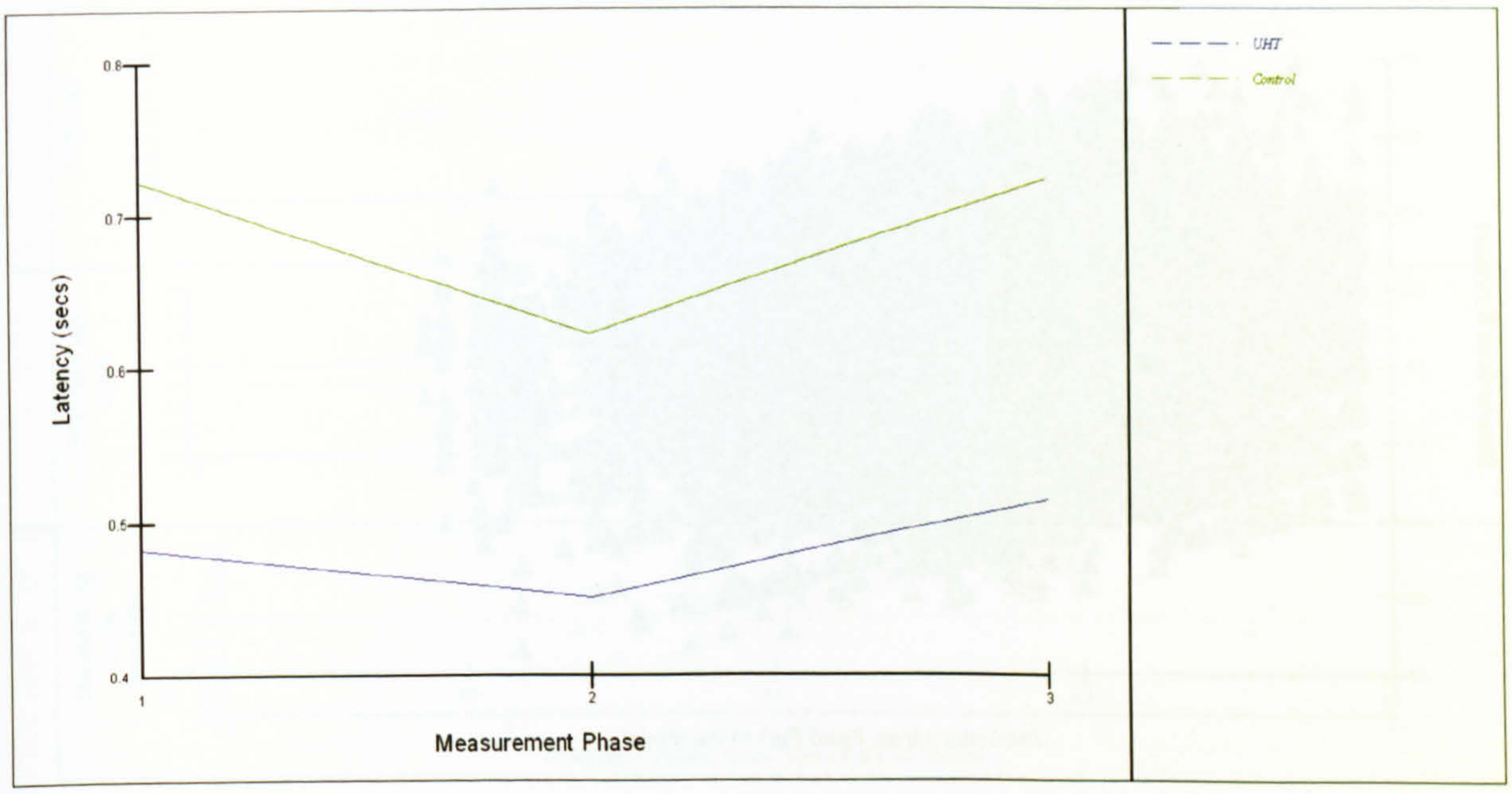

Figure 7.36 The predicted effect of *prior treatment* on the *latency* to press either lever, across different levels of *measurement phase* (from a prediction equation derived from the model specified in Equation 7.80).

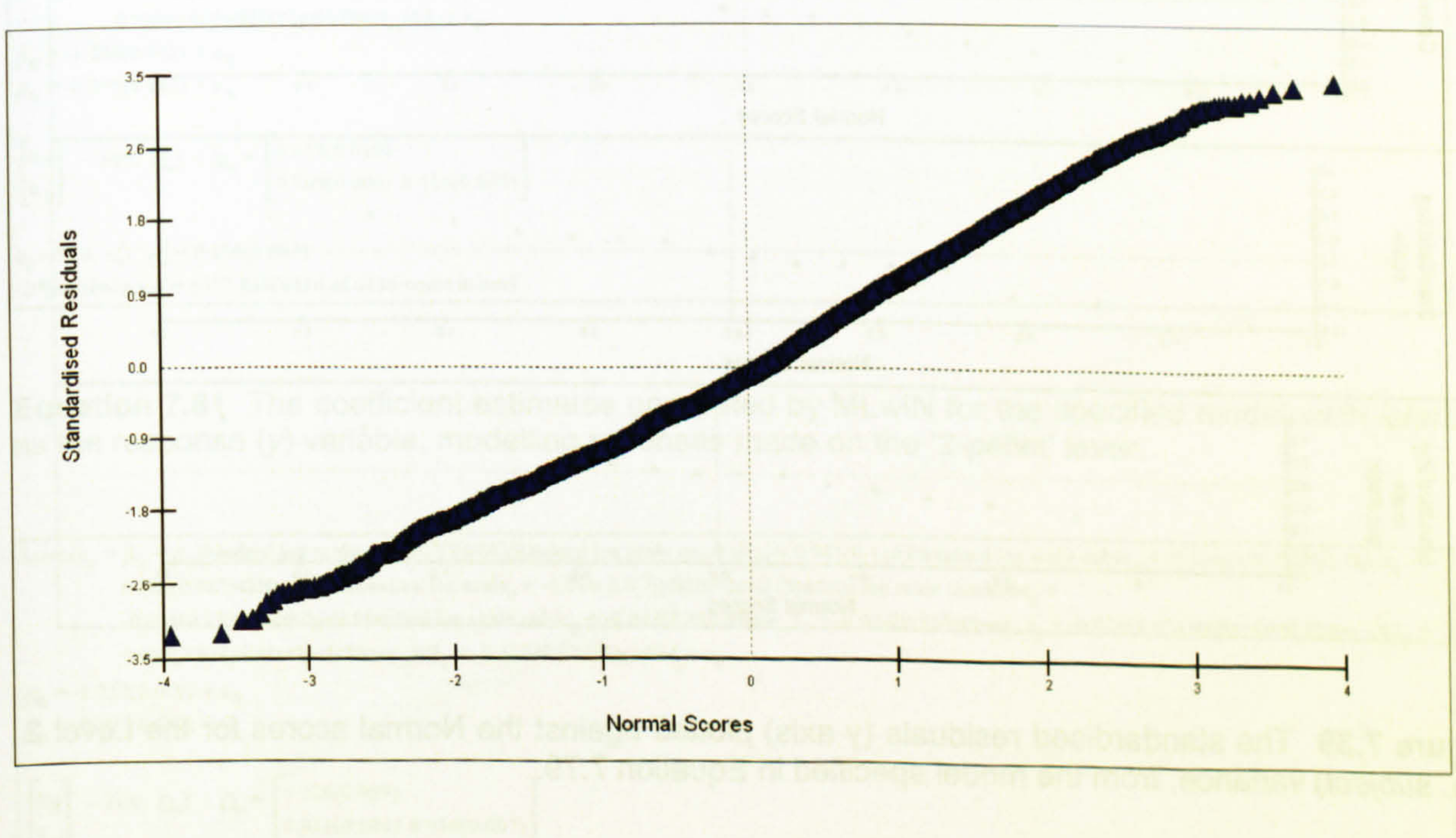

Figure 7.37 The standardised residuals (y-axis) plotted against the Normal scores for the Level 1 (i.e. *trial*) variance, from the model specified in Equation 7.79.

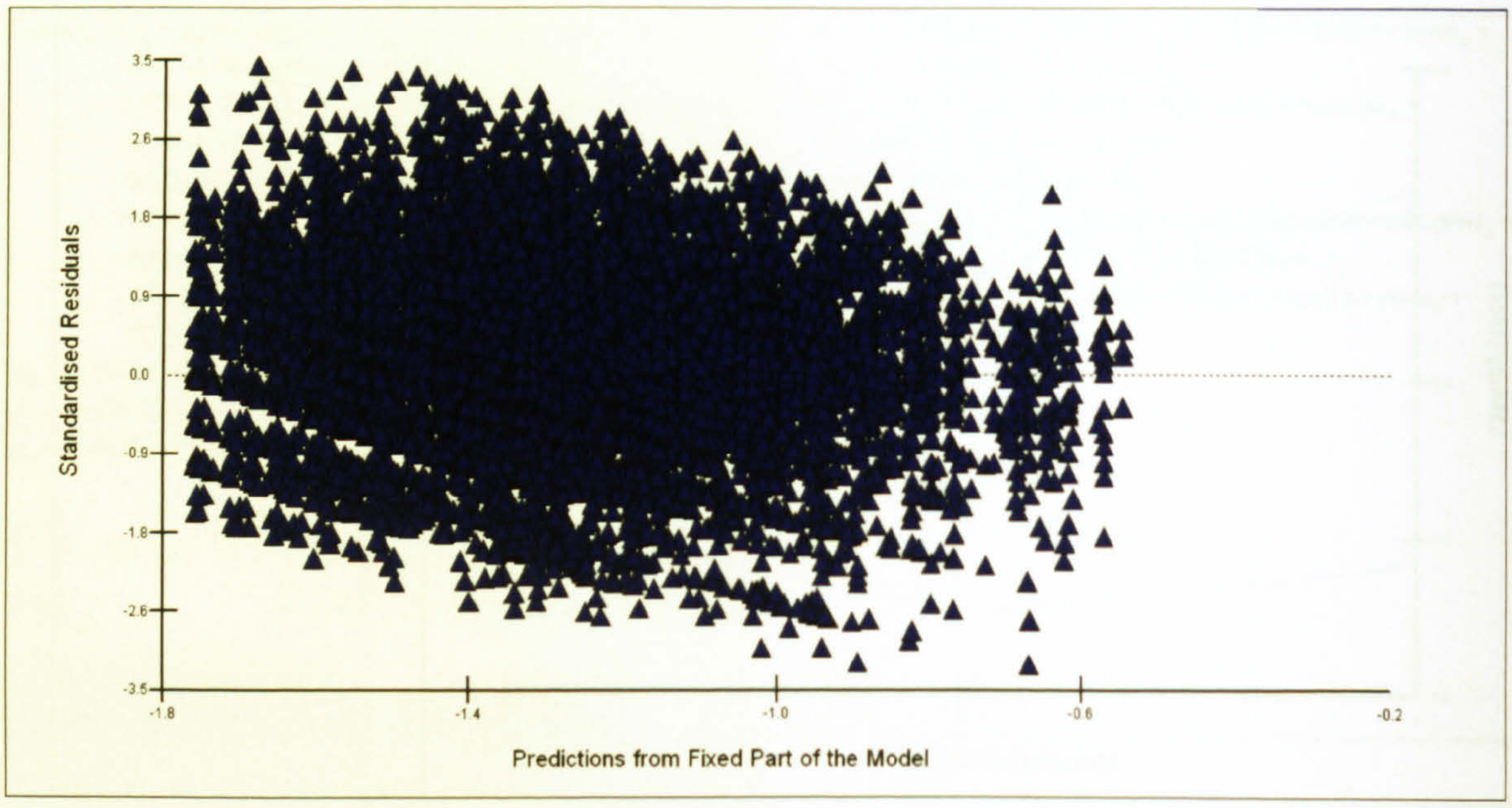

Figure 7.38 The standardised residuals (y-axis) plotted against the predicted values from the fixed part of the model (prior to backtransformation) for the Level 1 (i.e. *trial*) variance, from the model specified in Equation 7.79.

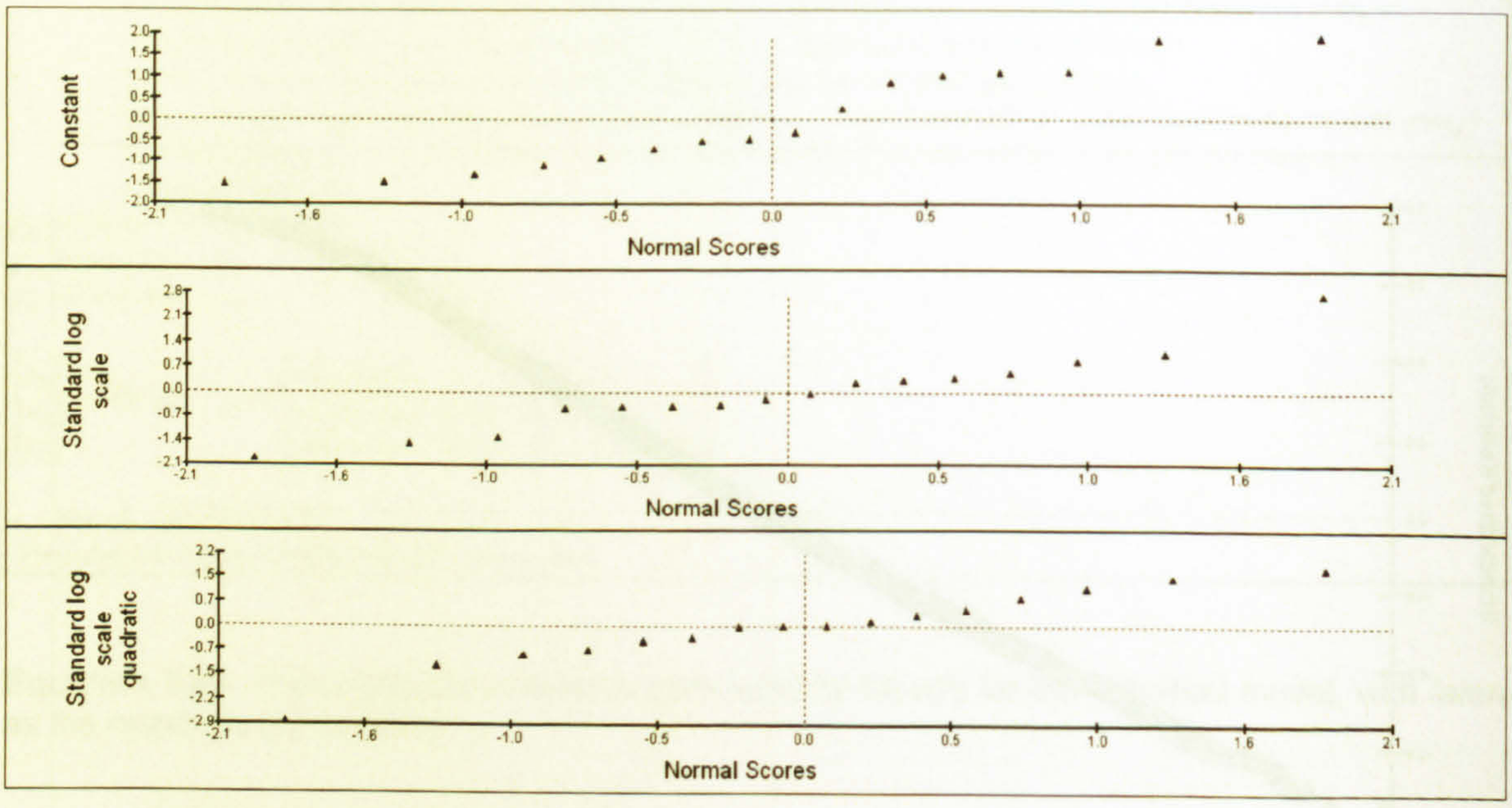

Figure 7.39 The standardised residuals (y-axis) plotted against the Normal scores for the Level 2 (i.e. *subject*) variance, from the model specified in Equation 7.79.

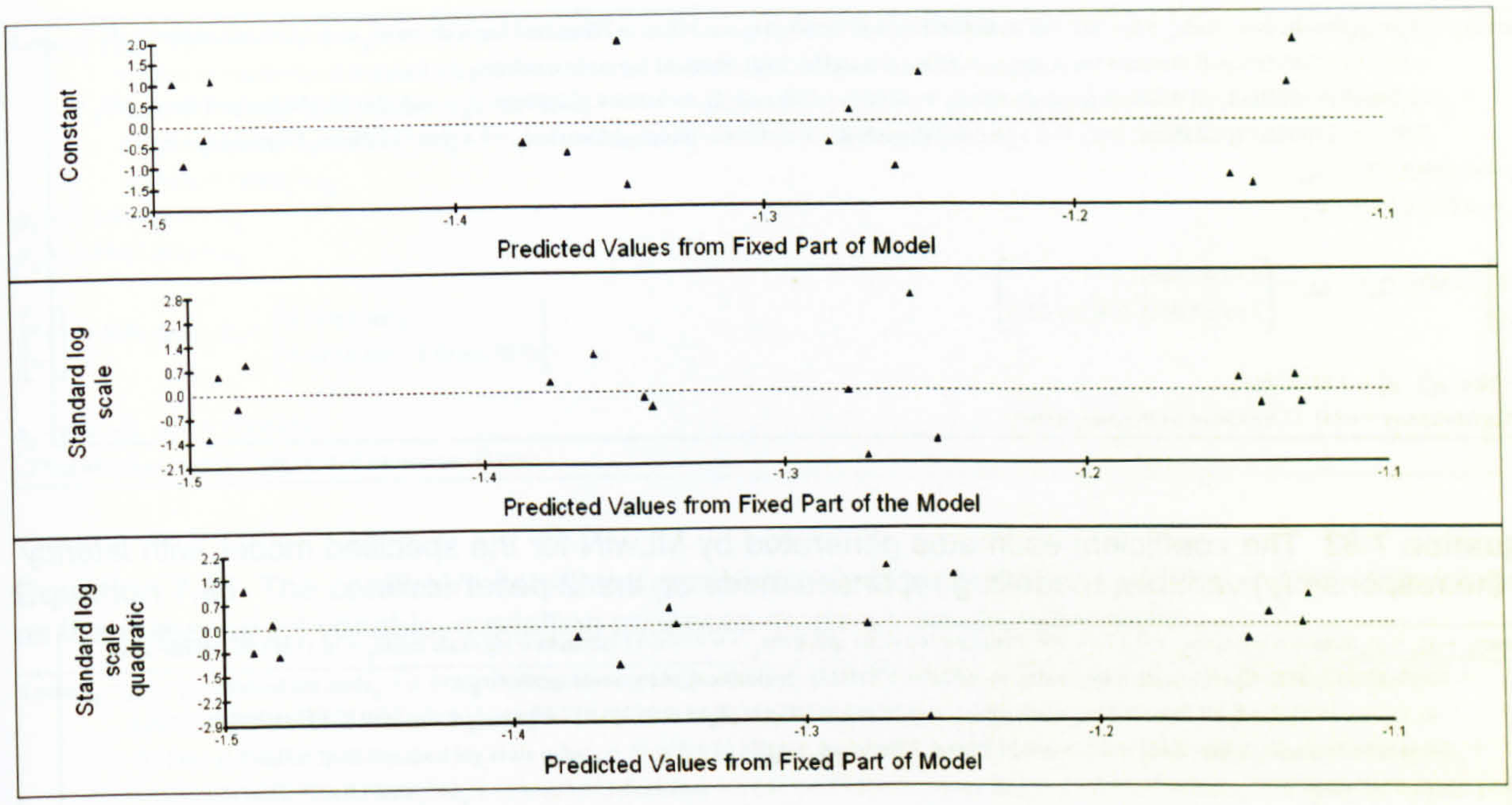

Figure 7.40 The standardised residuals (y-axis) plotted against the predicted values from the fixed part of the model (prior to backtransformation) for the Level 2 (i.e. *subject*) variance, from the model specified in Equation 7.79.

$$\begin{aligned}\text{Latency}_{ij} = {} & \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + 0.468(0.066)\text{Standard log scale quadratic}_{ij} + 0.983(0.118)\text{Standard log scale cubic}_{ij} + -0.126(0.086)\text{4kHz=2pell}_{j} + \\ & 0.509(0.087)\text{4kHz=2pell.Standard log scale}_{ij} + -0.019(0.079)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + \\ & -0.849(0.159)\text{4kHz=2pell.Standard log scale cubic}_{ij} + -0.043(0.019)\text{Phase\_2}_{ij} + -0.014(0.019)\text{Phase\_3}_{ij} + -0.026(0.024)\text{4kHz=2pell.Phase\_2\#2}_{ij} + \\ & 0.059(0.024)\text{4kHz=2pell.Phase\_3\#2}_{ij} + e_{ij}\end{aligned}$$

$$\beta_{0j} = -1.286(0.061) + u_{0j}$$

$$\beta_{1j} = -0.895(0.062) + u_{1j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.028(0.010) & \\ 0.009(0.006) & 0.016(0.007) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.155(0.003)$$

$-2*loglikelihood = 6377.846$(6436 of 6436 cases in use)

Equation 7.81 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable, modelling reponses made on the '2-pellet' lever.

$$\begin{aligned}\text{Latency}_{ij} = {} & \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + 0.467(0.066)\text{Standard log scale quadratic}_{ij} + 0.985(0.118)\text{Standard log scale cubic}_{ij} + -0.126(0.083)\text{4kHz=2pell}_{j} + \\ & 0.509(0.087)\text{4kHz=2pell.Standard log scale}_{ij} + -0.018(0.079)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + \\ & -0.851(0.159)\text{4kHz=2pell.Standard log scale cubic}_{ij} + -0.043(0.019)\text{Phase\_2}_{ij} + -0.014(0.019)\text{Phase\_3}_{ij} + -0.026(0.024)\text{4kHz=2pell.Phase\_2\#2}_{ij} + \\ & 0.058(0.024)\text{4kHz=2pell.Phase\_3\#2}_{ij} + -0.117(0.073)\text{Enriched}_{j} + e_{ij}\end{aligned}$$

$$\beta_{0j} = -1.227(0.070) + u_{0j}$$

$$\beta_{1j} = -0.895(0.062) + u_{1j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.026(0.009) & \\ 0.011(0.006) & 0.016(0.007) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.155(0.003)$$

$-2*loglikelihood = 6375.550$(6436 of 6436 cases in use)

Equation 7.82 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable, modelling reponses made on the '2-pellet' lever.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + 0.470(0.066)\text{Standard log scale quadratic}_{ij} + 0.978(0.117)\text{Standard log scale cubic}_{ij} + -0.122(0.083)\text{4kHz=2pell}_{j} + 0.502(0.087)\text{4kHz=2pell.Standard log scale}_{ij} + -0.019(0.078)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + -0.846(0.159)\text{4kHz=2pell.Standard log scale cubic}_{ij} + -0.067(0.022)\text{Phase\_2}_{ij} + -0.086(0.022)\text{Phase\_3}_{ij} + -0.028(0.024)\text{4kHz=2pell.Phase\_2\#2}_{ij} + 0.057(0.024)\text{4kHz=2pell.Phase\_3\#2}_{ij} + -0.174(0.075)\text{Enriched}_{j} + 0.045(0.024)\text{Phase\_2.Enriched}_{ij} + 0.139(0.024)\text{Phase\_3.Enriched}_{ij} + e_{ij}$$

$$\beta_{0j} = -1.199(0.070) + u_{0j}$$

$$\beta_{1j} = -0.892(0.062) + u_{1j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.026(0.009) & \\ 0.010(0.006) & 0.016(0.007) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.155(0.003)$$

*-2*loglikelihood* = 6341.172(6436 of 6436 cases in use)

Equation 7.83 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable, modelling reponses made on the '2-pellet' lever.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + 0.472(0.066)\text{Standard log scale quadratic}_{ij} + 0.979(0.118)\text{Standard log scale cubic}_{ij} + -0.113(0.082)\text{4kHz=2pell}_{j} + 0.506(0.087)\text{4kHz=2pell.Standard log scale}_{ij} + -0.021(0.078)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + -0.849(0.159)\text{4kHz=2pell.Standard log scale cubic}_{ij} + -0.084(0.017)\text{Phase\_2}_{ij} + -0.053(0.017)\text{Phase\_3}_{ij} + -0.176(0.074)\text{Enriched}_{j} + 0.044(0.024)\text{Phase\_2.Enriched}_{ij} + 0.139(0.024)\text{Phase\_3.Enriched}_{ij} + e_{ij}$$

$$\beta_{0j} = -1.203(0.070) + u_{0j}$$

$$\beta_{1j} = -0.895(0.062) + u_{1j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.026(0.009) & \\ 0.010(0.006) & 0.016(0.007) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.155(0.003)$$

*-2*loglikelihood* = 6353.309(6436 of 6436 cases in use)

Equation 7.84 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable, modelling reponses made on the '2-pellet' lever.

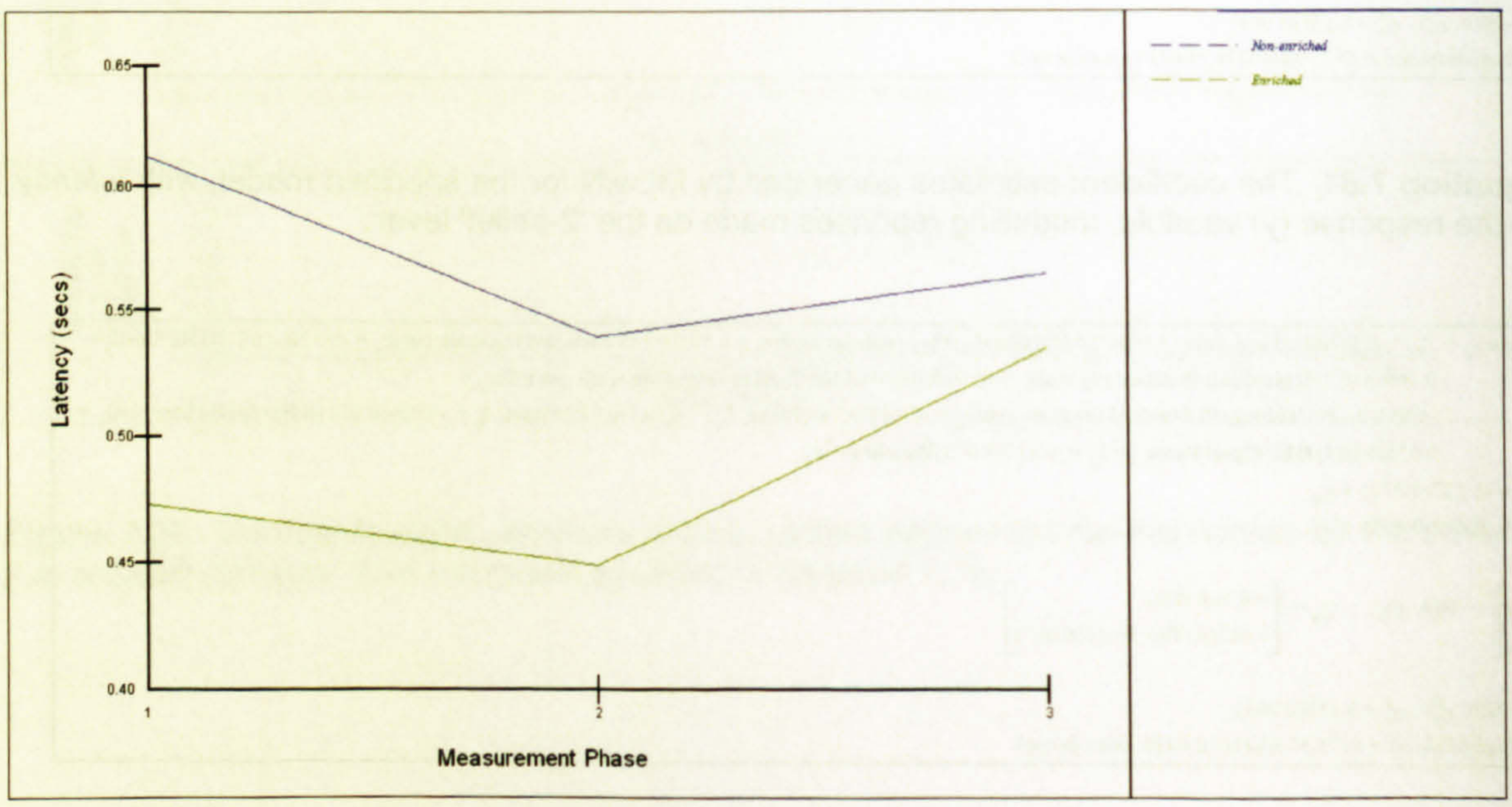

Figure 7.41 The predicted effect of *treatment* on the *latency* to press the '2-pellet' lever, across different levels of *measurement phase* (from a prediction equation derived from the model specified in Equation 7.84).

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + 0.473(0.066)\text{Standard log scale quadratic}_{ij} + 0.974(0.117)\text{Standard log scale cubic}_{ij} + -0.123(0.064)\text{4kHz=2pell}_{j} + 0.502(0.087)\text{4kHz=2pell.Standard log scale}_{ij} + -0.021(0.078)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + -0.841(0.159)\text{4kHz=2pell.Standard log scale cubic}_{ij} + -0.067(0.022)\text{Phase\_2}_{ij} + -0.087(0.022)\text{Phase\_3}_{ij} + -0.028(0.024)\text{4kHz=2pell.Phase\_2\#2}_{ij} + 0.057(0.024)\text{4kHz=2pell.Phase\_3\#2}_{ij} + -0.168(0.058)\text{Enriched}_{j} + 0.045(0.024)\text{Phase\_2.Enriched}_{ij} + 0.139(0.024)\text{Phase\_3.Enriched}_{ij} + -0.191(0.057)\text{UHT}_{j} + e_{ij}$$

$$\beta_{0j} = -1.106(0.061) + u_{0j}$$

$$\beta_{1j} = -0.894(0.062) + u_{1j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.015(0.005) & \\ 0.007(0.005) & 0.016(0.007) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.155(0.003)$$

$-2*\textit{loglikelihood} = 6332.986(6436 \text{ of } 6436 \text{ cases in use})$

Equation 7.85 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable, modelling reponses made on the '2-pellet' lever.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + 0.392(0.074)\text{Standard log scale quadratic}_{ij} + 0.861(0.139)\text{Standard log scale cubic}_{ij} + -0.121(0.064)\text{4kHz=2pell}_{j} + 0.497(0.088)\text{4kHz=2pell.Standard log scale}_{ij} + -0.022(0.078)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + -0.824(0.159)\text{4kHz=2pell.Standard log scale cubic}_{ij} + -0.065(0.022)\text{Phase\_2}_{ij} + -0.085(0.022)\text{Phase\_3}_{ij} + -0.030(0.024)\text{4kHz=2pell.Phase\_2\#2}_{ij} + 0.056(0.024)\text{4kHz=2pell.Phase\_3\#2}_{ij} + -0.163(0.059)\text{Enriched}_{j} + 0.045(0.024)\text{Phase\_2.Enriched}_{ij} + 0.137(0.024)\text{Phase\_3.Enriched}_{ij} + -0.231(0.062)\text{UHT}_{j} + -0.160(0.085)\text{UHT.Standard log scale}_{ij} + 0.154(0.063)\text{UHT.Standard log scale quadratic}_{ij} + 0.186(0.133)\text{UHT.Standard log scale cubic}_{ij} + e_{ij}$$

$$\beta_{0j} = -1.089(0.063) + u_{0j}$$

$$\beta_{1j} = -0.804(0.077) + u_{1j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.015(0.005) & \\ 0.007(0.005) & 0.016(0.007) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.154(0.003)$$

$-2*\textit{loglikelihood} = 6321.554(6436 \text{ of } 6436 \text{ cases in use})$

Equation 7.86 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable, modelling reponses made on the '2-pellet' lever.

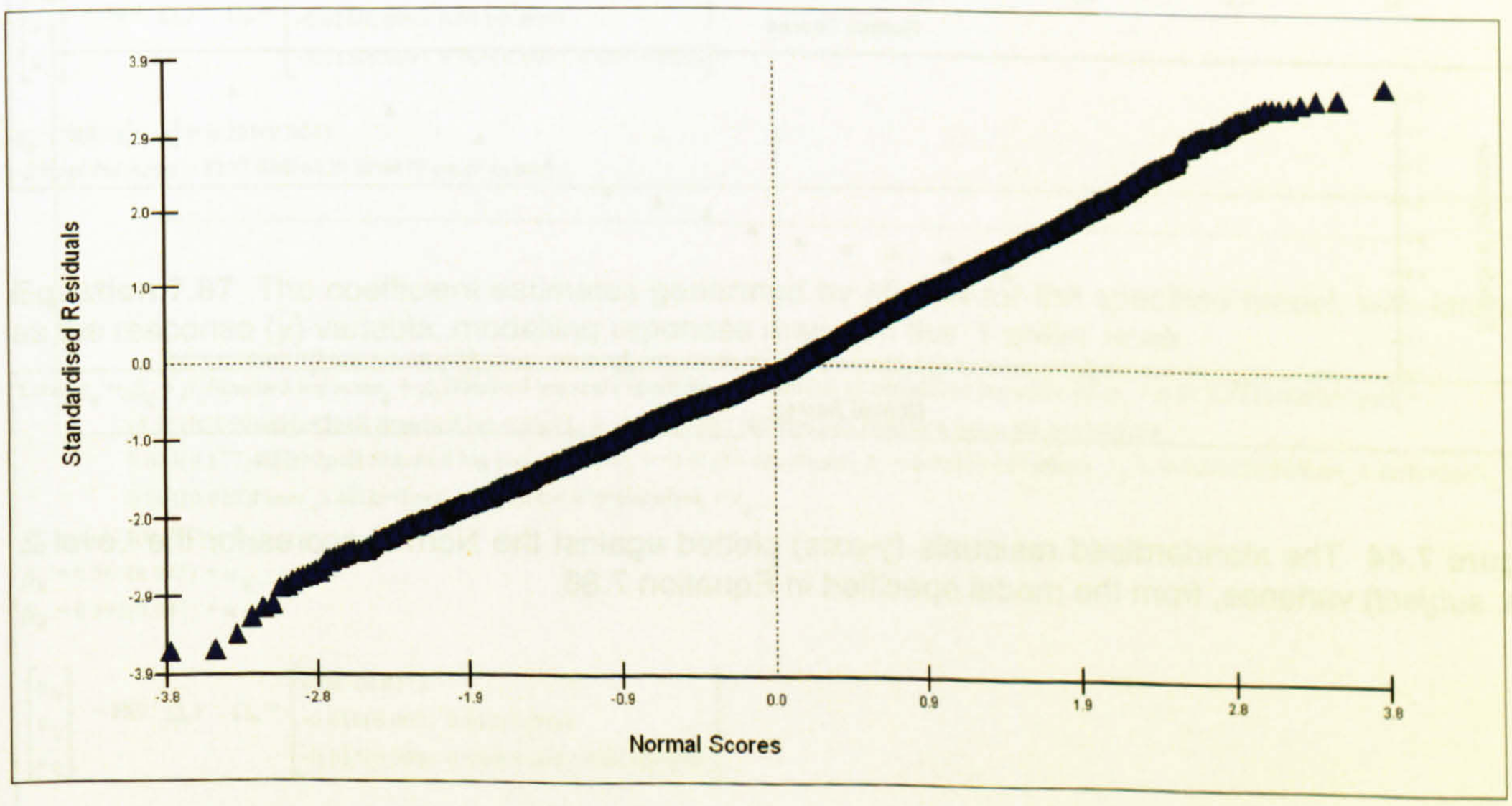

Figure 7.42 The standardised residuals (y-axis) plotted against the Normal scores for the Level 1 (i.e. *trial*) variance, from the model specified in Equation 7.86.

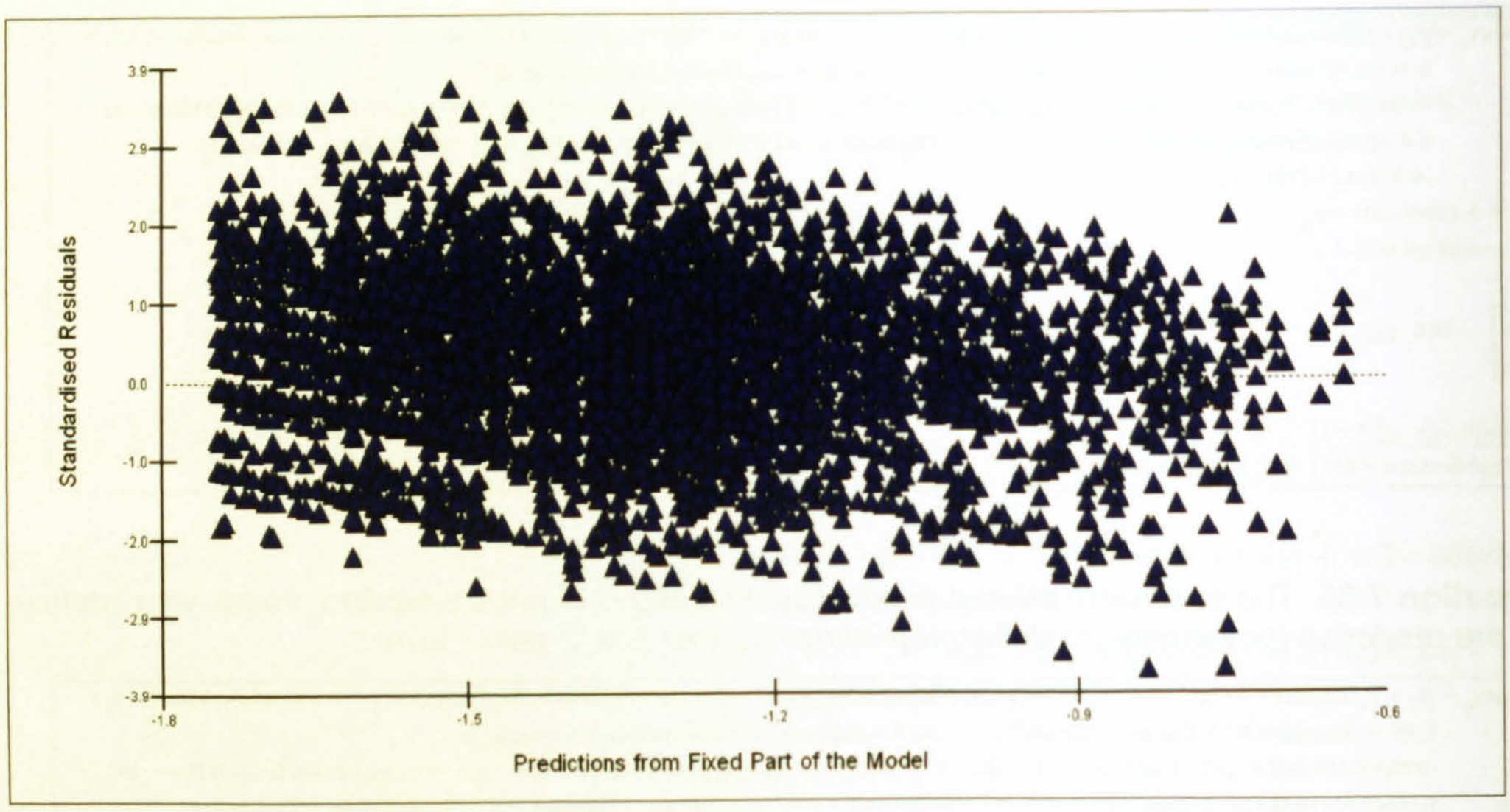

Figure 7.43 The standardised residuals (y-axis) plotted against the predicted values from the fixed part of the model (prior to backtransformation) for the Level 1 (i.e. *trial*) variance, from the model specified in Equation 7.86.

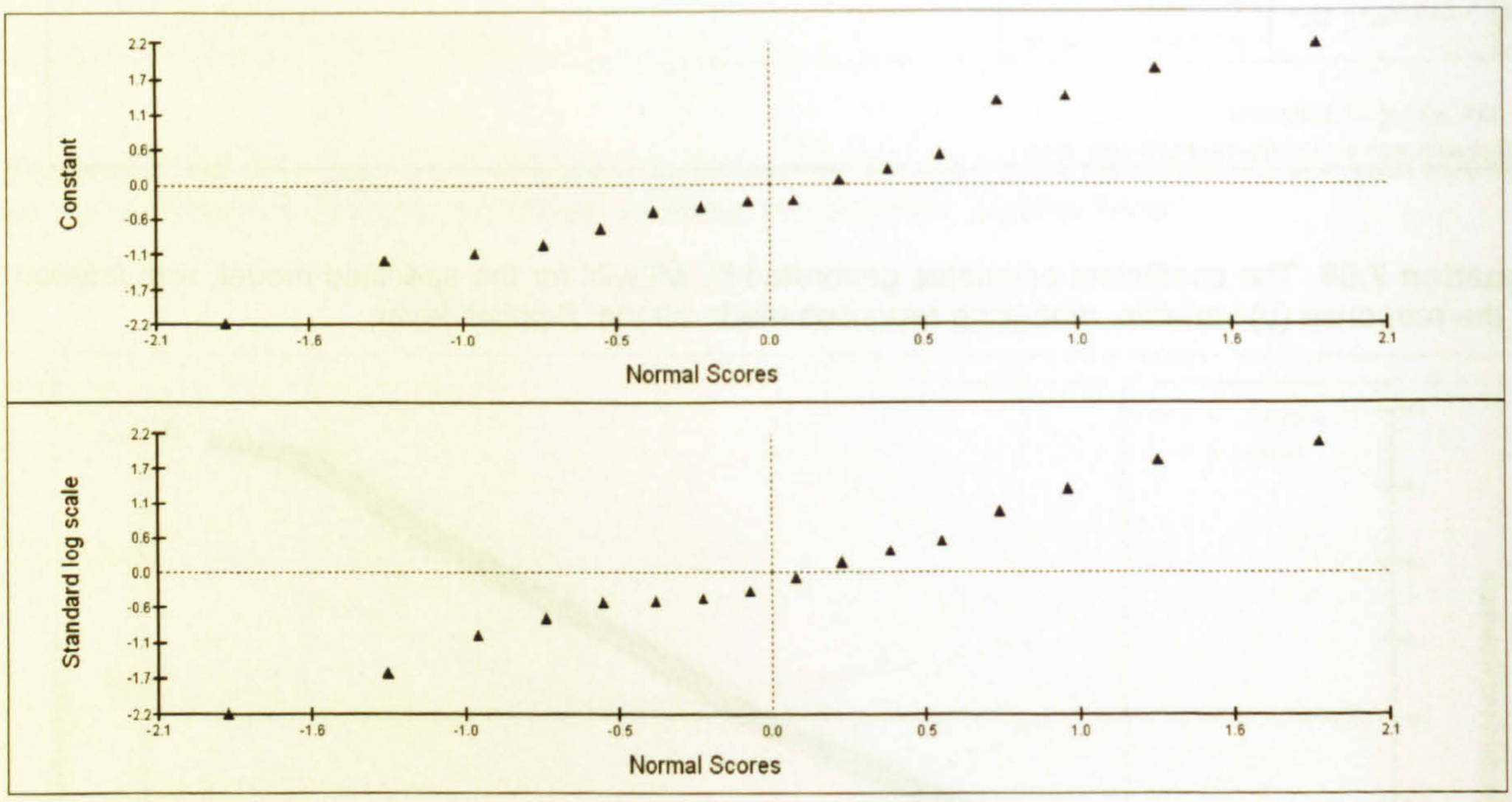

Figure 7.44 The standardised residuals (y-axis) plotted against the Normal scores for the Level 2 (i.e. *subject*) variance, from the model specified in Equation 7.86.

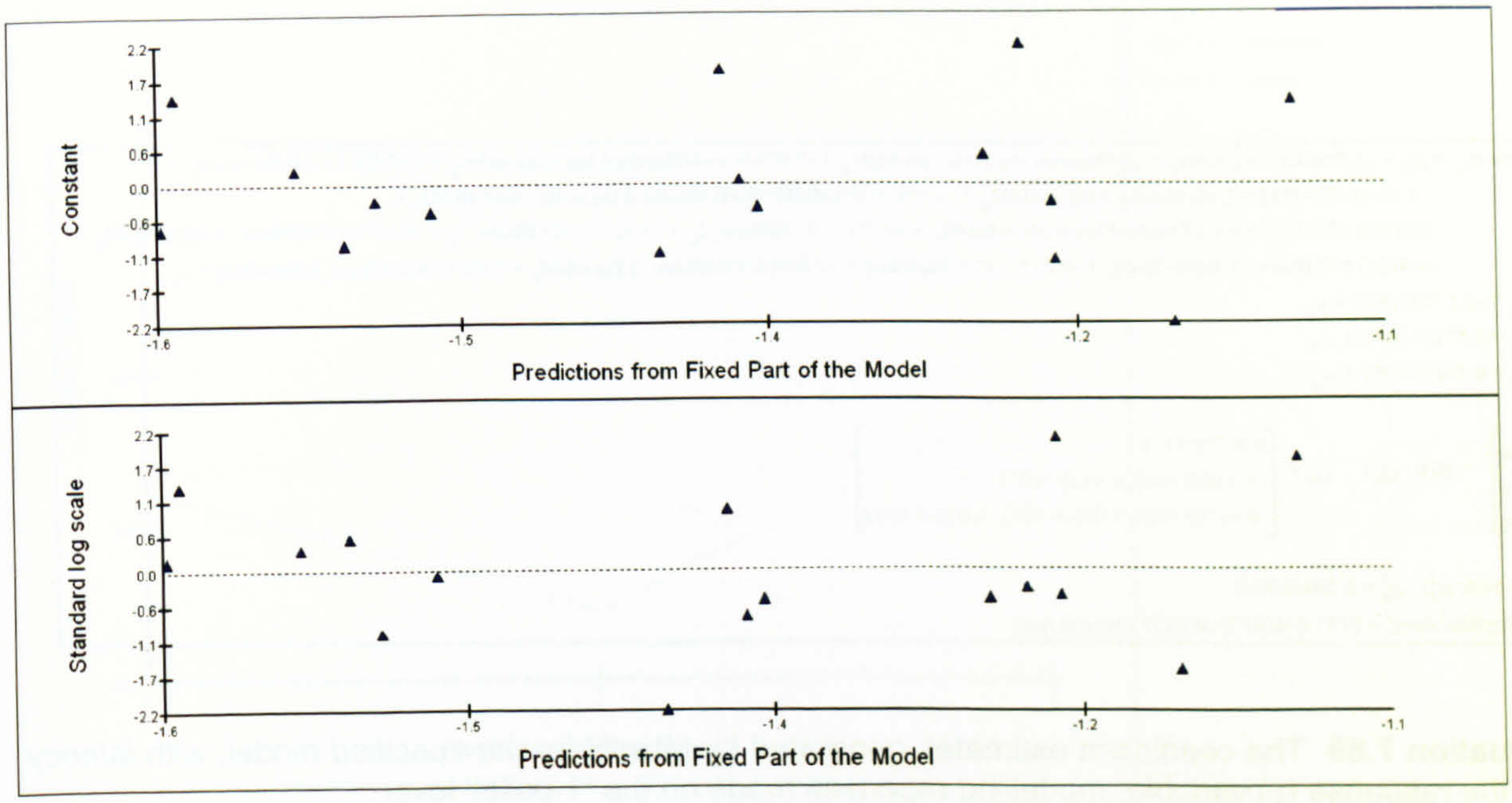

Figure 7.45 The standardised residuals (y-axis) plotted against the predicted values from the fixed part of the model (prior to backtransformation) for the Level 2 (i.e. *subject*) variance, from the model specified in Equation 7.86.

$$\begin{aligned}
\text{Latency}_{ij} = {} & \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -0.608(0.122)\text{Standard log scale cubic}_{ij} + 0.056(0.120)\text{4kHz=2pell}_{j} + \\
& -0.352(0.090)\text{4kHz=2pell.Standard log scale\#1}_{ij} + -0.395(0.086)\text{4kHz=2pell.Standard log scale quadratic\#1}_{ij} + \\
& 0.057(0.177)\text{4kHz=2pell.Standard log scale cubic\#1}_{ij} + -0.059(0.018)\text{Phase_2}_{ij} + 0.018(0.018)\text{Phase_3}_{ij} + -0.023(0.028)\text{Phase_2.4kHz=2pell}_{ij} + \\
& 0.041(0.028)\text{Phase_3.4kHz=2pell}_{ij} + e_{ij} \\
\beta_{0j} = {} & -1.297(0.084) + u_{0j} \\
\beta_{1j} = {} & 0.567(0.065) + u_{1j} \\
\beta_{2j} = {} & 0.489(0.048) + u_{2j} \\
\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} & \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.055(0.020) & & \\ -0.014(0.008) & 0.012(0.005) & \\ -0.015(0.009) & 0.006(0.005) & 0.001(0.006) \end{bmatrix} \\
e_{ij} \sim \text{N}(0,\ \sigma_e^2) & \quad \sigma_e^2 = 0.205(0.004) \\
\textit{-2*loglikelihood} = {} & 8207.634(6479 \text{ of } 6479 \text{ cases in use})
\end{aligned}$$

Equation 7.87 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable, modelling reponses made on the '1-pellet' lever.

$$\begin{aligned}
\text{Latency}_{ij} = {} & \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -0.601(0.122)\text{Standard log scale cubic}_{ij} + 0.057(0.111)\text{4kHz=2pell}_{j} + \\
& -0.351(0.090)\text{4kHz=2pell.Standard log scale\#1}_{ij} + -0.406(0.085)\text{4kHz=2pell.Standard log scale quadratic\#1}_{ij} + \\
& 0.043(0.177)\text{4kHz=2pell.Standard log scale cubic\#1}_{ij} + -0.059(0.018)\text{Phase_2}_{ij} + 0.018(0.018)\text{Phase_3}_{ij} + -0.024(0.028)\text{Phase_2.4kHz=2pell}_{ij} + \\
& 0.041(0.028)\text{Phase_3.4kHz=2pell}_{ij} + -0.244(0.075)\text{Enriched}_{j} + e_{ij} \\
\beta_{0j} = {} & -1.175(0.087) + u_{0j} \\
\beta_{1j} = {} & 0.567(0.065) + u_{1j} \\
\beta_{2j} = {} & 0.493(0.047) + u_{2j} \\
\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} & \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.047(0.017) & & \\ -0.018(0.008) & 0.012(0.006) & \\ -0.015(0.008) & 0.006(0.004) & 0.001(0.006) \end{bmatrix} \\
e_{ij} \sim \text{N}(0,\ \sigma_e^2) & \quad \sigma_e^2 = 0.205(0.004) \\
\textit{-2*loglikelihood} = {} & 8199.477(6479 \text{ of } 6479 \text{ cases in use})
\end{aligned}$$

Equation 7.88 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable, modelling reponses made on the '1-pellet' lever.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + \text{-0.608(0.121)Standard log scale cubic}_{ij} + \text{0.055(0.111)4kHz=2pell}_{j} +$$
$$\text{-0.355(0.090)4kHz=2pell.Standard log scale\#2}_{ij} + \text{-0.398(0.086)4kHz=2pell.Standard log scale quadratic\#2}_{ij} +$$
$$\text{0.060(0.176)4kHz=2pell.Standard log scale cubic\#2}_{ij} + \text{-0.011(0.023)Phase\_2}_{ij} + \text{-0.024(0.023)Phase\_3}_{ij} + \text{-0.027(0.028)Phase\_2.4kHz=2pell}_{ij} +$$
$$\text{0.042(0.028)Phase\_3.4kHz=2pell}_{ij} + \text{-0.235(0.076)Enriched}_{j} + \text{-0.095(0.028)Phase\_2.Enriched}_{ij} + \text{0.085(0.028)Phase\_3.Enriched}_{ij} + e_{ij}$$

$$\beta_{0j} = \text{-1.179(0.087)} + u_{0j}$$
$$\beta_{1j} = \text{0.571(0.065)} + u_{1j}$$
$$\beta_{2j} = \text{0.494(0.048)} + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} \text{0.047(0.017)} & & \\ \text{-0.018(0.008)} & \text{0.012(0.005)} & \\ \text{-0.015(0.008)} & \text{0.006(0.004)} & \text{0.002(0.006)} \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = \text{0.204(0.004)}$$

*-2*loglikelihood* = 8155.634(6479 of 6479 cases in use)

Equation 7.89 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable, modelling reponses made on the '1-pellet' lever.

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + \text{-0.612(0.121)Standard log scale cubic}_{ij} + \text{0.061(0.110)4kHz=2pell}_{j} +$$
$$\text{-0.356(0.090)4kHz=2pell.Standard log scale\#1}_{ij} + \text{-0.402(0.085)4kHz=2pell.Standard log scale quadratic\#1}_{ij} +$$
$$\text{0.059(0.176)4kHz=2pell.Standard log scale cubic\#1}_{ij} + \text{-0.023(0.019)Phase\_2}_{ij} + \text{-0.006(0.019)Phase\_3}_{ij} + \text{-0.235(0.076)Enriched}_{j} +$$
$$\text{-0.094(0.028)Phase\_2.Enriched}_{ij} + \text{0.085(0.028)Phase\_3.Enriched}_{ij} + e_{ij}$$

$$\beta_{0j} = \text{-1.181(0.087)} + u_{0j}$$
$$\beta_{1j} = \text{0.573(0.065)} + u_{1j}$$
$$\beta_{2j} = \text{0.494(0.048)} + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} \text{0.047(0.017)} & & \\ \text{-0.018(0.008)} & \text{0.012(0.006)} & \\ \text{-0.015(0.008)} & \text{0.006(0.004)} & \text{0.001(0.006)} \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = \text{0.204(0.004)}$$

*-2*loglikelihood* = 8161.994(6479 of 6479 cases in use)

Equation 7.90 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable, modelling reponses made on the '1-pellet' lever.

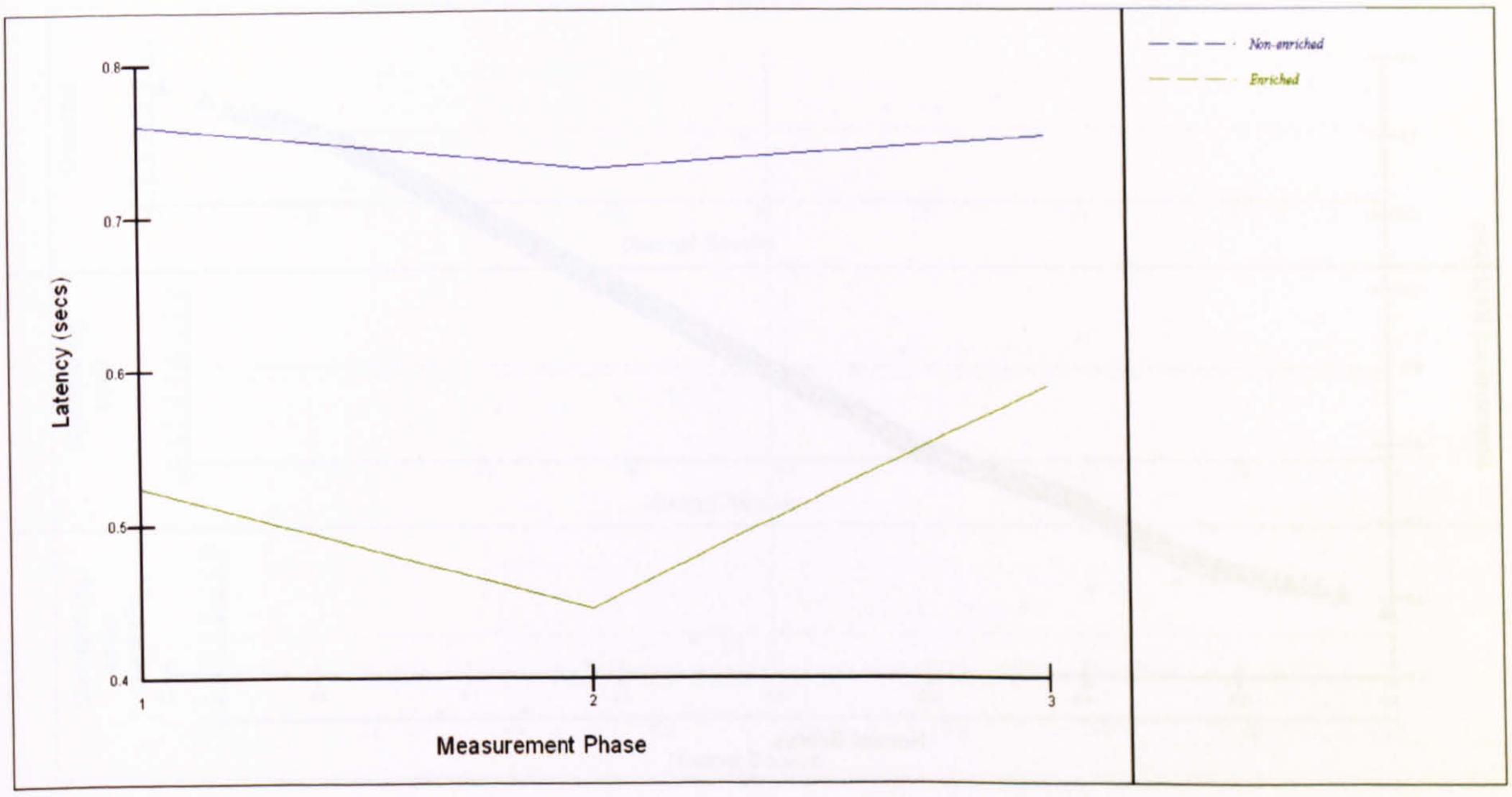

Figure 7.46 The predicted effect of *treatment* on the *latency* to press the '1-pellet' lever, across different levels of *measurement phase* (from a prediction equation derived from the model specified in Equation 7.90).

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -0.613(0.122)\text{Standard log scale cubic}_{ij} + 0.055(0.094)\text{4kHz=2pell}_{j} + -0.356(0.089)\text{4kHz=2pell.Standard log scale\#2}_{ij} + -0.393(0.086)\text{4kHz=2pell.Standard log scale quadratic\#2}_{ij} + 0.068(0.176)\text{4kHz=2pell.Standard log scale cubic\#2}_{ij} + -0.011(0.023)\text{Phase\_2}_{ij} + -0.024(0.023)\text{Phase\_3}_{ij} + -0.028(0.028)\text{4kHz=2pell.Phase\_2}_{ij} + 0.042(0.028)\text{4kHz=2pell.Phase\_3}_{ij} + -0.227(0.069)\text{Enriched}_{j} + -0.095(0.028)\text{Phase\_2.Enriched}_{ij} + 0.085(0.028)\text{Phase\_3.Enriched}_{ij} + -0.146(0.067)\text{UHT}_{j} + e_{ij}$$

$$\beta_{0j} = -1.110(0.081) + u_{0j}$$

$$\beta_{1j} = 0.570(0.064) + u_{1j}$$

$$\beta_{2j} = 0.488(0.048) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.033(0.012) & & \\ -0.014(0.007) & 0.011(0.005) & \\ -0.011(0.007) & 0.005(0.004) & 0.002(0.006) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.204(0.004)$$

$$-2 * \textit{loglikelihood} = 8152.285(6479 \text{ of } 6479 \text{ cases in use})$$

Equation 7.91 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (y) variable, modelling reponses made on the '1-pellet' lever.

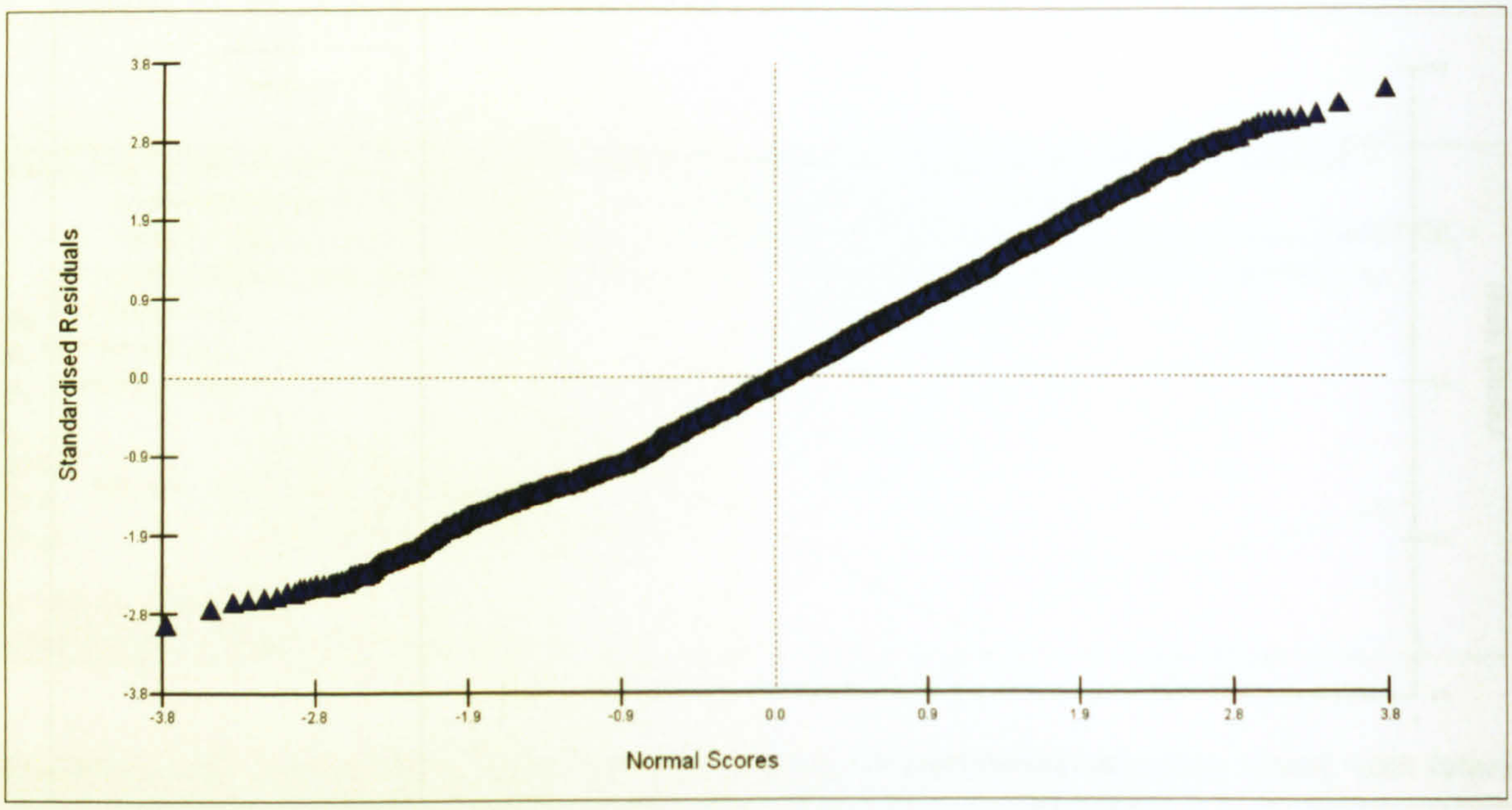

Figure 7.47 The standardised residuals (y-axis) plotted against the Normal scores for the Level 1 (i.e. *trial*) variance, from the model specified in Equation 7.89.

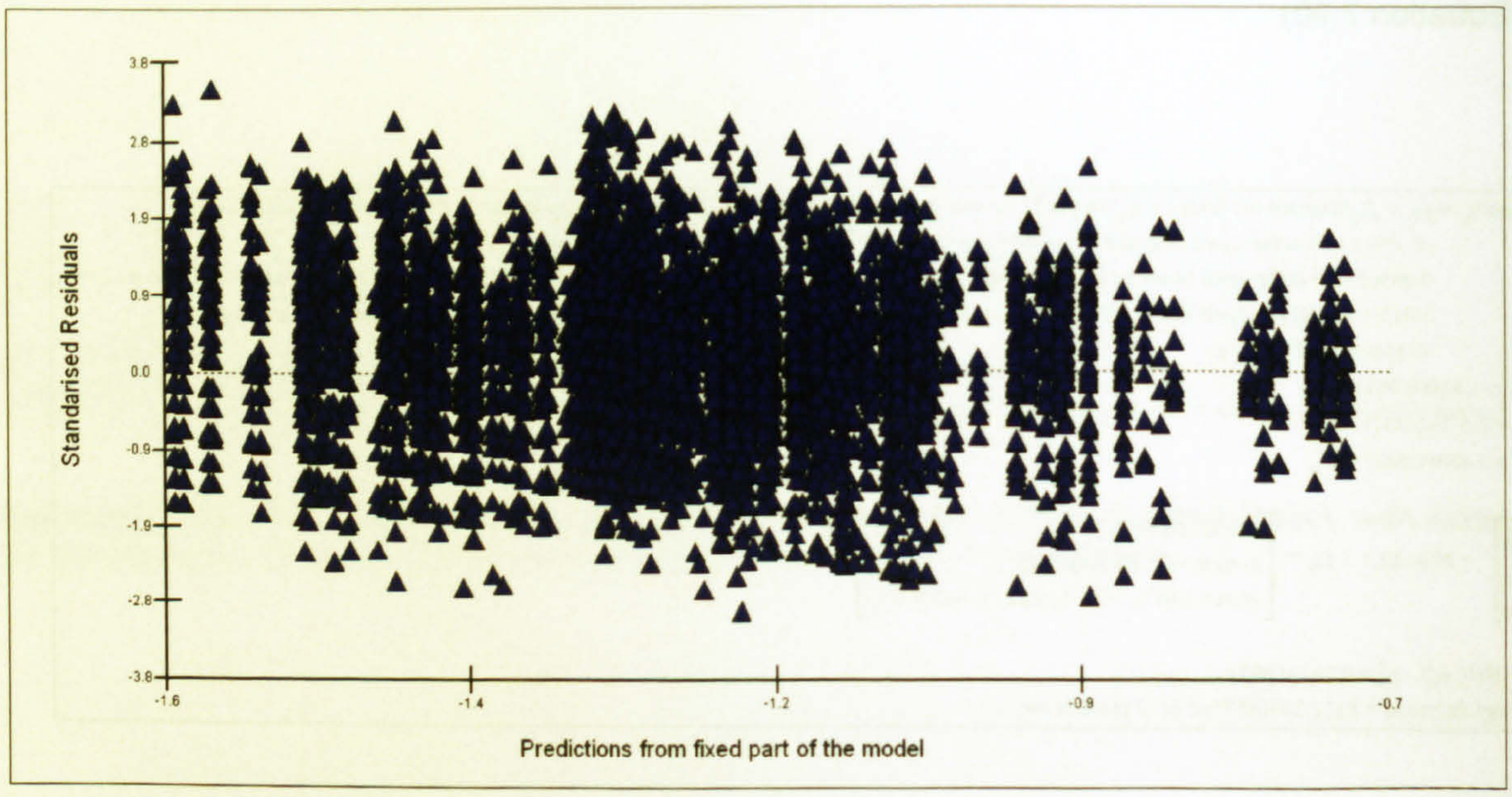

Figure 7.48 The standardised residuals (y-axis) plotted against the predicted values from the fixed part of the model (prior to backtransformation) for the Level 1 (i.e. *trial*) variance, from the model specified in Equation 7.89.

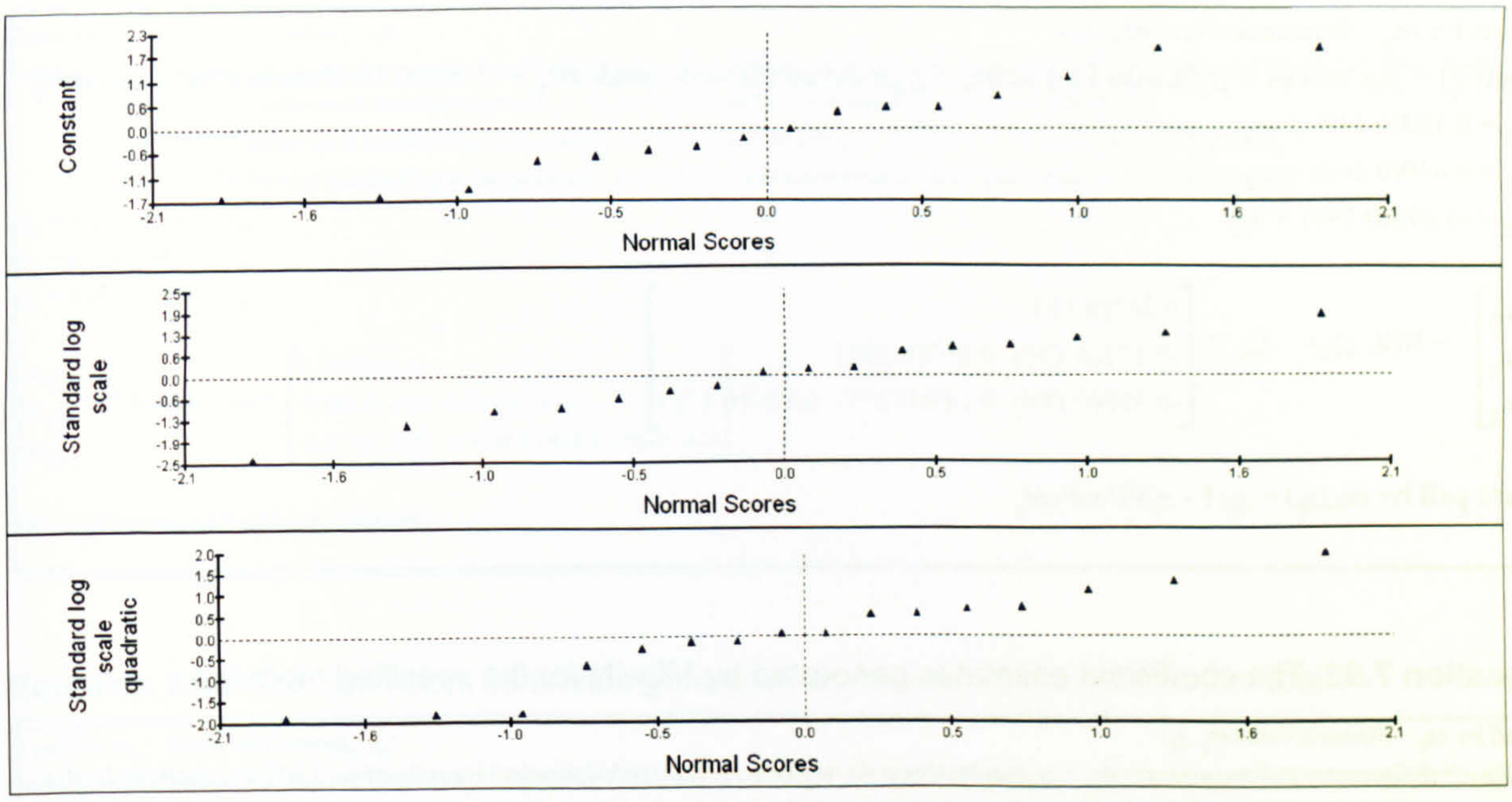

Figure 7.49 The standardised residuals (y-axis) plotted against the Normal scores for the Level 2 (i.e. *subject*) variance, from the model specified in Equation 7.89.

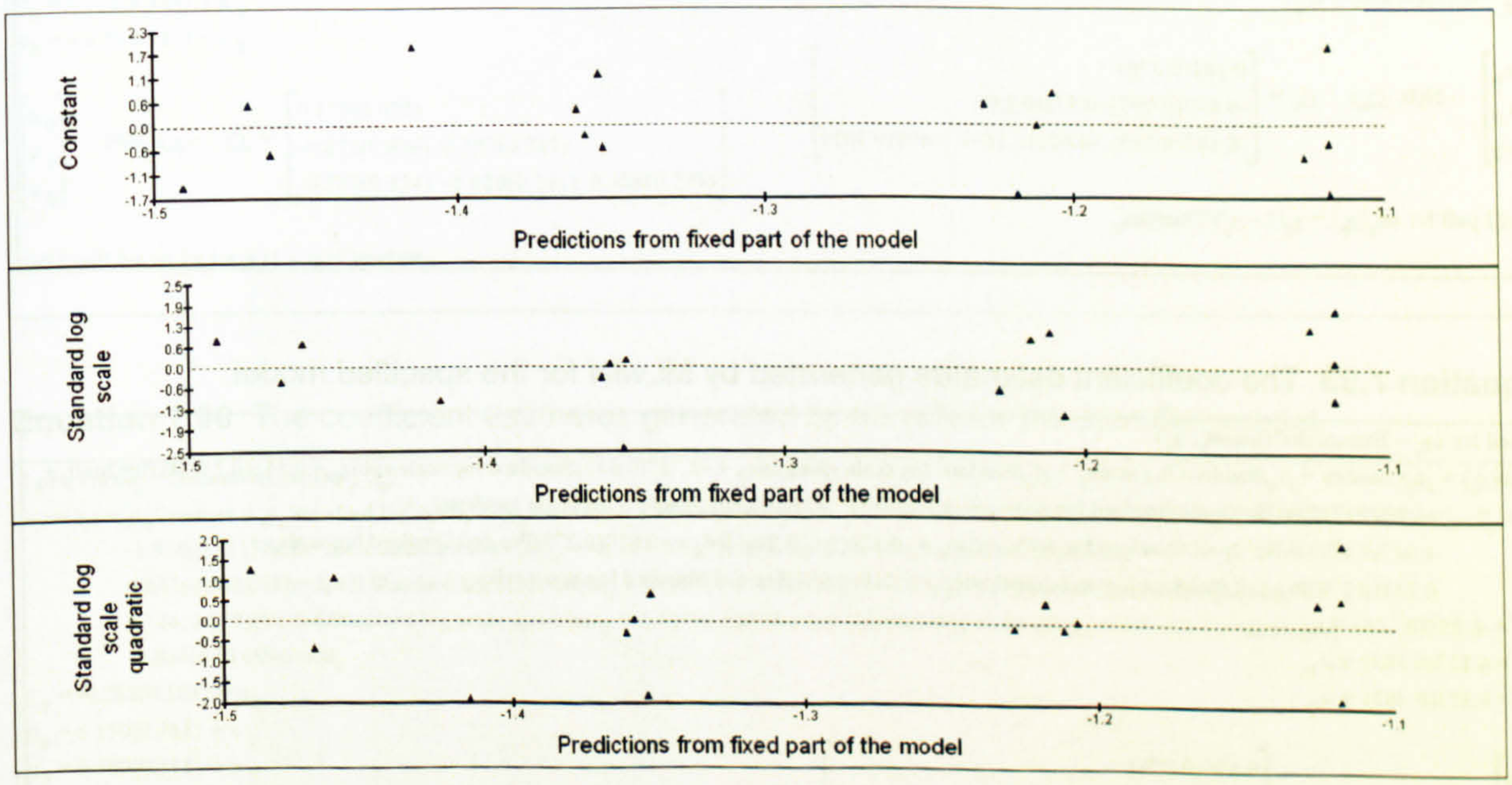

Figure 7.50 The standardised residuals (y-axis) plotted against the predicted values from the fixed part of the model (prior to backtransformation) for the Level 2 (i.e. *subject*) variance, from the model specified in Equation 7.89.

2 pell lvr on$_{ij}$ ~ Binomial(Constant$_{ij}$, π_{ij})

logit(π_{ij}) = β_{0j}Constant + β_{1j}Standard log scale$_{ij}$ + β_{2j}Standard log scale quadratic$_{ij}$ + -5.926(0.366)Standard log scale cubic$_{ij}$

β_{0j} = 0.183(0.138) + u_{0j}

β_{1j} = 4.639(0.263) + u_{1j}

β_{2j} = -0.698(0.294) + u_{2j}

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) \; : \; \Omega_u = \begin{bmatrix} 0.265(0.105) & & \\ -0.171(0.129) & 0.697(0.282) & \\ -0.329(0.186) & 0.118(0.259) & 0.957(0.471) \end{bmatrix}$$

var(2 pell lvr on$_{ij}$|π_{ij}) = π_{ij}(1 - π_{ij})/Constant$_{ij}$

Equation 7.92 The coefficient estimates generated by MLwiN for the specified model.

2 pell lvr on$_{ij}$ ~ Binomial(Constant$_{ij}$, π_{ij})

logit(π_{ij}) = β_{0j}Constant + β_{1j}Standard log scale$_{ij}$ + β_{2j}Standard log scale quadratic$_{ij}$ + -8.135(0.563)Standard log scale cubic$_{ij}$ + 0.551(0.236)4kHz=2pell$_j$ + -1.994(0.504)4kHz=2pell.Standard log scale$_{ij}$ + -1.341(0.517)4kHz=2pell.Standard log scale quadratic$_{ij}$ + 4.034(0.800)4kHz=2pell.Standard log scale cubic$_{ij}$

β_{0j} = -0.115(0.161) + u_{0j}

β_{1j} = 5.668(0.352) + u_{1j}

β_{2j} = 0.452(0.356) + u_{2j}

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) \; : \; \Omega_u = \begin{bmatrix} 0.184(0.076) & & \\ -0.077(0.097) & 0.574(0.237) & \\ -0.145(0.119) & -0.023(0.189) & 0.494(0.301) \end{bmatrix}$$

var(2 pell lvr on$_{ij}$|π_{ij}) = π_{ij}(1 - π_{ij})/Constant$_{ij}$

Equation 7.93 The coefficient estimates generated by MLwiN for the specified model.

2 pell lvr on$_{ij}$ ~ Binomial(Constant$_{ij}$, π_{ij})

logit(π_{ij}) = β_{0j}Constant + β_{1j}Standard log scale$_{ij}$ + β_{2j}Standard log scale quadratic$_{ij}$ + -8.747(0.643)Standard log scale cubic$_{ij}$ + 0.551(0.237)4kHz=2pell$_j$ + -1.994(0.510)4kHz=2pell.Standard log scale$_{ij}$ + -1.333(0.517)4kHz=2pell.Standard log scale quadratic$_{ij}$ + 4.047(0.802)4kHz=2pell.Standard log scale cubic$_{ij}$ + -0.123(0.075)Pre-fed$_{ij}$ + -0.918(0.278)Pre-fed.Standard log scale$_{ij}$ + 0.141(0.278)Pre-fed.Standard log scale quadratic$_{ij}$ + 1.210(0.620)Pre-fed.Standard log scale cubic$_{ij}$

β_{0j} = -0.055(0.166) + u_{0j}

β_{1j} = 6.127(0.383) + u_{1j}

β_{2j} = 0.383(0.382) + u_{2j}

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) \; : \; \Omega_u = \begin{bmatrix} 0.186(0.076) & & \\ -0.073(0.098) & 0.594(0.244) & \\ -0.146(0.119) & -0.026(0.192) & 0.493(0.301) \end{bmatrix}$$

var(2 pell lvr on$_{ij}$|π_{ij}) = π_{ij}(1 - π_{ij})/Constant$_{ij}$

Equation 7.94 The coefficient estimates generated by MLwiN for the specified model.

$$\text{2 pell lvr on}_{ij} \sim \text{Binomial}(\text{Constant}_{ij}, \pi_{ij})$$

$$\begin{aligned}\text{logit}(\pi_{ij}) = {} & \beta_{0j}\text{Constant} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -8.772(0.643)\text{Standard log scale cubic}_{ij} + 0.565(0.245)\text{4kHz=2pell}_{j} + \\ & -1.995(0.510)\text{4kHz=2pell.Standard log scale}_{ij} + -1.345(0.518)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + \\ & 4.052(0.802)\text{4kHz=2pell.Standard log scale cubic}_{ij} + -0.123(0.075)\text{Pre-fed}_{ij} + -0.918(0.278)\text{Pre-fed.Standard log scale}_{ij} + \\ & 0.143(0.278)\text{Pre-fed.Standard log scale quadratic}_{ij} + 1.211(0.620)\text{Pre-fed.Standard log scale cubic}_{ij} + 0.159(0.197)\text{UHT}_{j}\end{aligned}$$

$$\beta_{0j} = -0.134(0.198) + u_{0j}$$

$$\beta_{1j} = 6.139(0.383) + u_{1j}$$

$$\beta_{2j} = 0.387(0.382) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0, \Omega_u) \; : \; \Omega_u = \begin{bmatrix} 0.200(0.081) & & \\ -0.084(0.102) & 0.590(0.243) & \\ -0.175(0.126) & -0.024(0.191) & 0.496(0.302) \end{bmatrix}$$

$$\text{var}(\text{2 pell lvr on}_{ij} | \pi_{ij}) = \pi_{ij}(1 - \pi_{ij})/\text{Constant}_{ij}$$

Equation 7.95 The coefficient estimates generated by MLwiN for the specified model.

$$\text{2 pell lvr on}_{ij} \sim \text{Binomial}(\text{Constant}_{ij}, \pi_{ij})$$

$$\begin{aligned}\text{logit}(\pi_{ij}) = {} & \beta_{0j}\text{Constant} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -8.765(0.643)\text{Standard log scale cubic}_{ij} + 0.583(0.231)\text{4kHz=2pell}_{j} + \\ & -1.996(0.511)\text{4kHz=2pell.Standard log scale}_{ij} + -1.340(0.516)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + \\ & 4.043(0.802)\text{4kHz=2pell.Standard log scale cubic}_{ij} + -0.123(0.075)\text{Pre-fed}_{ij} + -0.920(0.278)\text{Pre-fed.Standard log scale}_{ij} + \\ & 0.144(0.278)\text{Pre-fed.Standard log scale quadratic}_{ij} + 1.216(0.619)\text{Pre-fed.Standard log scale cubic}_{ij} + 0.437(0.168)\text{Enriched}_{j}\end{aligned}$$

$$\beta_{0j} = -0.274(0.182) + u_{0j}$$

$$\beta_{1j} = 6.136(0.384) + u_{1j}$$

$$\beta_{2j} = 0.390(0.381) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0, \Omega_u) \; : \; \Omega_u = \begin{bmatrix} 0.175(0.072) & & \\ -0.071(0.096) & 0.597(0.245) & \\ -0.203(0.124) & -0.020(0.191) & 0.489(0.299) \end{bmatrix}$$

$$\text{var}(\text{2 pell lvr on}_{ij} | \pi_{ij}) = \pi_{ij}(1 - \pi_{ij})/\text{Constant}_{ij}$$

Equation 7.96 The coefficient estimates generated by MLwiN for the specified model.

$$\text{2 pell lvr on}_{ij} \sim \text{Binomial}(\text{Constant}_{ij}, \pi_{ij})$$

$$\begin{aligned}\text{logit}(\pi_{ij}) = {} & \beta_{0j}\text{Constant} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -8.776(0.644)\text{Standard log scale cubic}_{ij} + 0.577(0.232)\text{4kHz=2pell}_{j} + \\ & -1.996(0.511)\text{4kHz=2pell.Standard log scale}_{ij} + -1.344(0.517)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + \\ & 4.042(0.802)\text{4kHz=2pell.Standard log scale cubic}_{ij} + -0.117(0.076)\text{Pre-fed}_{ij} + -0.926(0.278)\text{Pre-fed.Standard log scale}_{ij} + \\ & 0.146(0.278)\text{Pre-fed.Standard log scale quadratic}_{ij} + 1.217(0.620)\text{Pre-fed.Standard log scale cubic}_{ij} + 0.453(0.166)\text{Enriched}_{j} + \\ & -0.064(0.016)\text{Session}_{ij}\end{aligned}$$

$$\beta_{0j} = -0.282(0.183) + u_{0j}$$

$$\beta_{1j} = 6.150(0.384) + u_{1j}$$

$$\beta_{2j} = 0.389(0.381) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0, \Omega_u) \; : \; \Omega_u = \begin{bmatrix} 0.177(0.073) & & \\ -0.073(0.096) & 0.598(0.245) & \\ -0.209(0.126) & -0.020(0.192) & 0.492(0.301) \end{bmatrix}$$

$$\text{var}(\text{2 pell lvr on}_{ij} | \pi_{ij}) = \pi_{ij}(1 - \pi_{ij})/\text{Constant}_{ij}$$

Equation 7.97 The coefficient estimates generated by MLwiN for the specified model.

Observed	Expected (predicted from model)	$(O-E)^2$	$(O-E)^2$ / E
78	75.461781	6.44	0.09
142	131.34294	113.57	0.86
208	190.13885	319.02	1.68
236	257.09289	444.91	1.73
327	346.99666	399.87	1.15
451	455.90866	24.09	0.05
525	530.5563	30.87	0.06
594	587.45413	42.85	0.07
632	629.57427	5.88	0.01
684	671.32476	160.66	0.24
		Chi^2 statistic:	**5.943461**
		d.f.	**10**
		p	**0.82**

Table 7.6 Calculations pertaining to the Hosmer-Lemeshow goodness-of-fit test of the model specified in Equation 7.97; the 'Observed' and 'Expected' refer to the number of presses on the lever associated with 2 pellets of food, and the subgroups are ordered with respect to the 'Expected' values (with subgroup 1 having the lowest expected values, and subgroup 10 the greatest, with a roughly equal number of *trials* in each subgroup).

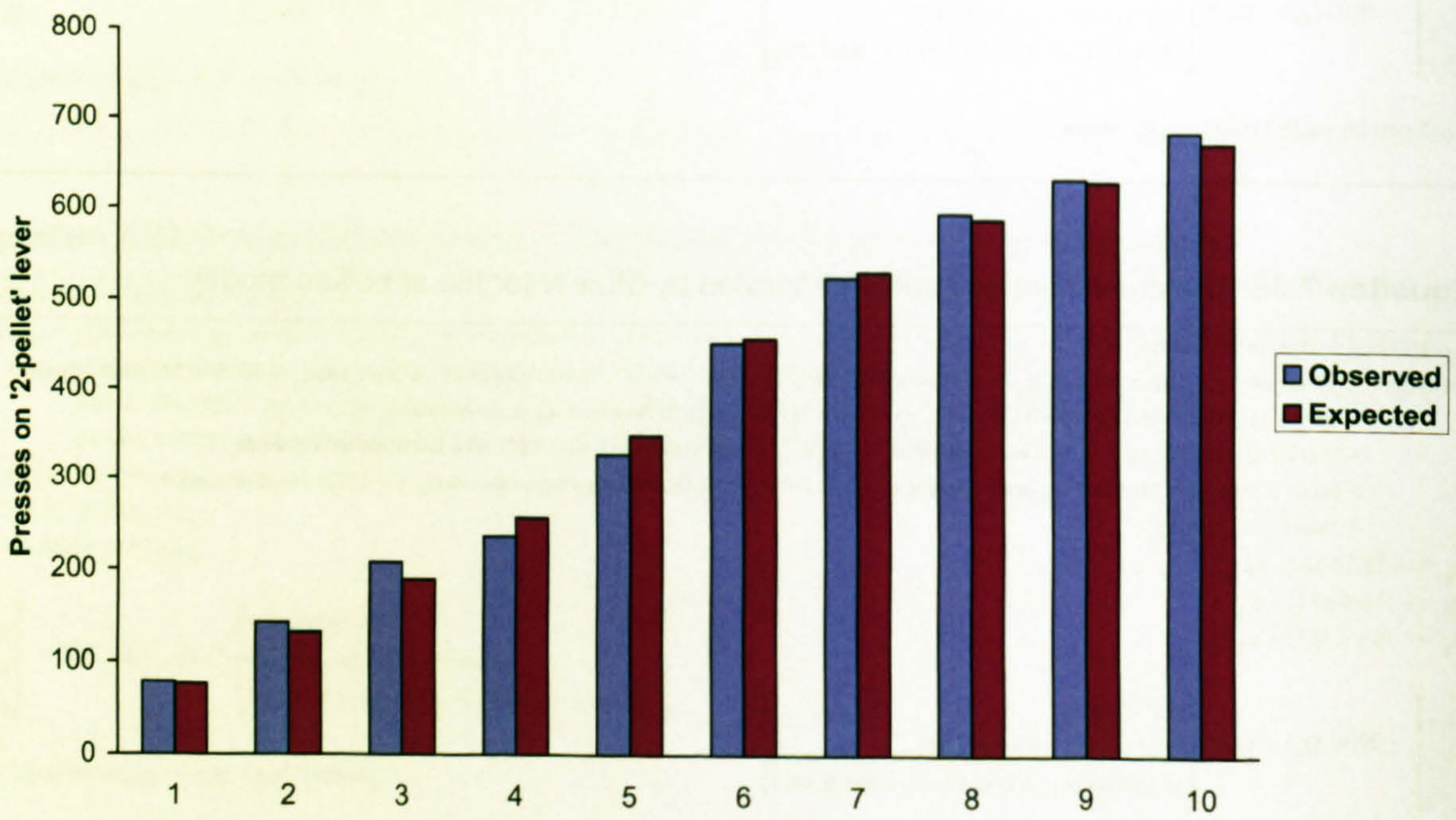

Figure 7.51 Plot of the 'Observed' and 'Expected' number of presses on the lever associated with 2 pellets of food, across the ordered subgroups (see Table 7.6).

$$\text{Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + 0.345(0.072)\text{Standard log scale cubic}_{ij} + e_{ij}$$

$$\beta_{0j} = -1.255(0.056) + u_{0j}$$

$$\beta_{1j} = -0.237(0.045) + u_{1j}$$

$$\beta_{2j} = 0.371(0.051) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.047(0.017) & & \\ -0.002(0.008) & 0.018(0.007) & \\ -0.007(0.011) & -0.008(0.007) & 0.025(0.014) \end{bmatrix}$$

$$e_{ij} \sim N(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.219(0.004)$$

$-2 * loglikelihood = 10098.300(7565 \text{ of } 7565 \text{ cases in use})$

Equation 7.98 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable.

$$\text{RecipRoot\_Rounded\_Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -0.646(0.096)\text{Standard log scale cubic}_{ij} + -0.059(0.017)\text{2 pell lvr on\_1}_{ij} + -1.249(0.057)\text{2 pell lvr on\_1.Standard log scale}_{ij} + 0.134(0.060)\text{2 pell lvr on\_1.Standard log scale quadratic}_{ij} + 1.494(0.129)\text{2 pell lvr on\_1.Standard log scale cubic}_{ij} + e_{ij}$$

$$\beta_{0j} = -1.167(0.054) + u_{0j}$$

$$\beta_{1j} = 0.491(0.050) + u_{1j}$$

$$\beta_{2j} = 0.423(0.053) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.042(0.016) & & \\ -0.005(0.006) & 0.010(0.004) & \\ -0.010(0.010) & -0.004(0.005) & 0.017(0.011) \end{bmatrix}$$

$$e_{ij} \sim N(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.200(0.003)$$

$-2 * loglikelihood = 9409.253(7565 \text{ of } 7565 \text{ cases in use})$

Equation 7.99 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable.

$$\text{RecipRoot\_Rounded\_Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -0.501(0.156)\text{Standard log scale cubic}_{ij} + 0.014(0.024)\text{2 pell lvr on\_1}_{ij} + -1.424(0.090)\text{2 pell lvr on\_1.Standard log scale}_{ij} + 0.058(0.107)\text{2 pell lvr on\_1.Standard log scale quadratic}_{ij} + 1.629(0.226)\text{2 pell lvr on\_1.Standard log scale cubic}_{ij} + 0.150(0.107)\text{4kHz=2pell}_{j} + -0.090(0.102)\text{4kHz=2pell.Standard log scale}_{ij} + -0.296(0.126)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + -0.461(0.232)\text{4kHz=2pell.Standard log scale cubic}_{ij} + -0.194(0.035)\text{4kHz=2pell.2 pell lvr on\_1}_{ij} + 0.422(0.126)\text{4kHz=2pell.2 pell lvr on\_1.Standard log scale}_{ij} + 0.388(0.155)\text{4kHz=2pell.2 pell lvr on\_1.Standard log scale quadratic}_{ij} + -0.249(0.325)\text{4kHz=2pell.2 pell lvr on\_1.Standard log scale cubic}_{ij} + e_{ij}$$

$$\beta_{0j} = -1.216(0.072) + u_{0j}$$

$$\beta_{1j} = 0.521(0.074) + u_{1j}$$

$$\beta_{2j} = 0.425(0.071) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.040(0.015) & & \\ -0.002(0.005) & 0.007(0.003) & \\ -0.010(0.009) & -0.003(0.004) & 0.014(0.010) \end{bmatrix}$$

$$e_{ij} \sim N(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.198(0.003)$$

$-2 * loglikelihood = 9330.703(7565 \text{ of } 7565 \text{ cases in use})$

Equation 7.100 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable.

$$\begin{aligned}
\text{RecipRoot_Rounded_Latency}_{ij} = {} & \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -0.470(0.154)\text{Standard log scale cubic}_{ij} + \\
& 0.019(0.024)\text{2 pell lvr on_1}_{ij} + -1.398(0.089)\text{2 pell lvr on_1.Standard log scale}_{ij} + \\
& 0.046(0.105)\text{2 pell lvr on_1.Standard log scale quadratic}_{ij} + 1.608(0.224)\text{2 pell lvr on_1.Standard log scale cubic}_{ij} + \\
& 0.149(0.108)\text{4kHz=2pell}_{j} + -0.089(0.101)\text{4kHz=2pell.Standard log scale}_{ij} + \\
& -0.288(0.125)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + -0.463(0.230)\text{4kHz=2pell.Standard log scale cubic}_{ij} + \\
& -0.193(0.035)\text{4kHz=2pell.2 pell lvr on_1}_{ij} + 0.425(0.125)\text{4kHz=2pell.2 pell lvr on_1.Standard log scale}_{ij} + \\
& 0.388(0.154)\text{4kHz=2pell.2 pell lvr on_1.Standard log scale quadratic}_{ij} + \\
& -0.280(0.321)\text{4kHz=2pell.2 pell lvr on_1.Standard log scale cubic}_{ij} + 0.133(0.010)\text{Pre-fed}_{ij} + e_{ij}
\end{aligned}$$

$$\beta_{0j} = -1.283(0.074) + u_{0j}$$
$$\beta_{1j} = 0.502(0.073) + u_{1j}$$
$$\beta_{2j} = 0.424(0.071) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.041(0.015) & & \\ -0.001(0.005) & 0.007(0.003) & \\ -0.010(0.009) & -0.003(0.004) & 0.014(0.009) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.194(0.003)$$

$-2*loglikelihood = 9162.965(7565 \text{ of } 7565 \text{ cases in use})$

Equation 7.101 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable.

$$\begin{aligned}
\text{RecipRoot_Rounded_Latency}_{ij} = {} & \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -0.478(0.154)\text{Standard log scale cubic}_{ij} + \\
& 0.018(0.024)\text{2 pell lvr on_1}_{ij} + -1.404(0.089)\text{2 pell lvr on_1.Standard log scale}_{ij} + \\
& 0.049(0.105)\text{2 pell lvr on_1.Standard log scale quadratic}_{ij} + 1.613(0.224)\text{2 pell lvr on_1.Standard log scale cubic}_{ij} + \\
& 0.115(0.108)\text{4kHz=2pell}_{j} + -0.099(0.101)\text{4kHz=2pell.Standard log scale}_{ij} + \\
& -0.287(0.125)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + -0.449(0.230)\text{4kHz=2pell.Standard log scale cubic}_{ij} + \\
& -0.191(0.035)\text{2 pell lvr on_1.4kHz=2pell}_{ij} + 0.440(0.125)\text{2 pell lvr on_1.4kHz=2pell.Standard log scale}_{ij} + \\
& 0.382(0.154)\text{2 pell lvr on_1.4kHz=2pell.Standard log scale quadratic}_{ij} + \\
& -0.299(0.321)\text{2 pell lvr on_1.4kHz=2pell.Standard log scale cubic}_{ij} + 0.101(0.014)\text{Pre-fed}_{ij} + \\
& 0.067(0.020)\text{4kHz=2pell.Pre-fed}_{ij} + e_{ij}
\end{aligned}$$

$$\beta_{0j} = -1.267(0.074) + u_{0j}$$
$$\beta_{1j} = 0.506(0.073) + u_{1j}$$
$$\beta_{2j} = 0.424(0.071) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.041(0.015) & & \\ -0.001(0.005) & 0.007(0.003) & \\ -0.010(0.009) & -0.003(0.004) & 0.014(0.010) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.194(0.003)$$

$-2*loglikelihood = 9152.180(7565 \text{ of } 7565 \text{ cases in use})$

Equation 7.102 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable.

$$\begin{aligned}
\text{RecipRoot_Rounded_Latency}_{ij} = {} & \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -0.479(0.154)\text{Standard log scale cubic}_{ij} + \\
& 0.007(0.027)\text{2 pell lvr on_1}_{ij} + -1.404(0.089)\text{2 pell lvr on_1.Standard log scale}_{ij} + \\
& 0.048(0.105)\text{2 pell lvr on_1.Standard log scale quadratic}_{ij} + 1.616(0.224)\text{2 pell lvr on_1.Standard log scale cubic}_{ij} + \\
& 0.085(0.109)\text{4kHz=2pell}_{j} + -0.106(0.101)\text{4kHz=2pell.Standard log scale}_{ij} + \\
& -0.284(0.125)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + -0.437(0.229)\text{4kHz=2pell.Standard log scale cubic}_{ij} + \\
& -0.139(0.040)\text{2 pell lvr on_1.4kHz=2pell}_{ij} + 0.443(0.125)\text{2 pell lvr on_1.4kHz=2pell.Standard log scale}_{ij} + \\
& 0.381(0.154)\text{2 pell lvr on_1.4kHz=2pell.Standard log scale quadratic}_{ij} + \\
& -0.305(0.321)\text{2 pell lvr on_1.4kHz=2pell.Standard log scale cubic}_{ij} + 0.091(0.019)\text{Pre-fed}_{ij} + \\
& 0.024(0.028)\text{2 pell lvr on_1.Pre-fed}_{ij} + 0.125(0.030)\text{4kHz=2pell.Pre-fed}_{ij} + \\
& -0.105(0.041)\text{2 pell lvr on_1.4kHz=2pell.Pre-fed}_{ij} + e_{ij}
\end{aligned}$$

$$\beta_{0j} = -1.262(0.074) + u_{0j}$$
$$\beta_{1j} = 0.508(0.073) + u_{1j}$$
$$\beta_{2j} = 0.424(0.071) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.041(0.015) & & \\ -0.002(0.005) & 0.007(0.003) & \\ -0.010(0.009) & -0.003(0.004) & 0.014(0.010) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.194(0.003)$$

*-2*loglikelihood* = 9144.339(7565 of 7565 cases in use)

Equation 7.103 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable.

$$\begin{aligned}
\text{RecipRoot_Rounded_Latency}_{ij} = {} & \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -0.478(0.154)\text{Standard log scale cubic}_{ij} + \\
& 0.007(0.027)\text{2 pell lvr on_1}_{ij} + -1.404(0.089)\text{2 pell lvr on_1.Standard log scale}_{ij} + \\
& 0.048(0.105)\text{2 pell lvr on_1.Standard log scale quadratic}_{ij} + 1.616(0.224)\text{2 pell lvr on_1.Standard log scale cubic}_{ij} + \\
& 0.075(0.103)\text{4kHz=2pell}_{j} + -0.106(0.101)\text{4kHz=2pell.Standard log scale}_{ij} + \\
& -0.285(0.125)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + -0.438(0.229)\text{4kHz=2pell.Standard log scale cubic}_{ij} + \\
& -0.139(0.040)\text{2 pell lvr on_1.4kHz=2pell}_{ij} + 0.443(0.125)\text{2 pell lvr on_1.4kHz=2pell.Standard log scale}_{ij} + \\
& 0.382(0.154)\text{2 pell lvr on_1.4kHz=2pell.Standard log scale quadratic}_{ij} + \\
& -0.304(0.321)\text{2 pell lvr on_1.4kHz=2pell.Standard log scale cubic}_{ij} + 0.091(0.019)\text{Pre-fed}_{ij} + \\
& 0.023(0.028)\text{2 pell lvr on_1.Pre-fed}_{ij} + 0.125(0.030)\text{4kHz=2pell.Pre-fed}_{ij} + \\
& -0.104(0.041)\text{2 pell lvr on_1.4kHz=2pell.Pre-fed}_{ij} + -0.146(0.090)\text{UHT}_{j} + e_{ij}
\end{aligned}$$

$$\beta_{0j} = -1.189(0.083) + u_{0j}$$
$$\beta_{1j} = 0.507(0.073) + u_{1j}$$
$$\beta_{2j} = 0.424(0.071) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.036(0.013) & & \\ 0.000(0.005) & 0.006(0.003) & \\ -0.011(0.009) & -0.003(0.004) & 0.014(0.009) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.194(0.003)$$

*-2*loglikelihood* = 9142.081(7565 of 7565 cases in use)

Equation 7.104 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable.

$$\begin{aligned}\text{RecipRoot\_Rounded\_Latency}_{ij} = {} & \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + \text{-0.471(0.154)Standard log scale cubic}_{ij} + \\ & \text{-0.049(0.029)2 pell lvr on\_1}_{ij} + \text{-1.402(0.088)2 pell lvr on\_1.Standard log scale}_{ij} + \\ & \text{0.055(0.105)2 pell lvr on\_1.Standard log scale quadratic}_{ij} + \text{1.611(0.223)2 pell lvr on\_1.Standard log scale cubic}_{ij} + \\ & \text{0.070(0.103)4kHz=2pell}_{j} + \text{-0.099(0.100)4kHz=2pell.Standard log scale}_{ij} + \\ & \text{-0.281(0.125)4kHz=2pell.Standard log scale quadratic}_{ij} + \text{-0.451(0.229)4kHz=2pell.Standard log scale cubic}_{ij} + \\ & \text{-0.130(0.040)2 pell lvr on\_1.4kHz=2pell}_{ij} + \text{0.443(0.125)2 pell lvr on\_1.4kHz=2pell.Standard log scale}_{ij} + \\ & \text{0.374(0.153)2 pell lvr on\_1.4kHz=2pell.Standard log scale quadratic}_{ij} + \\ & \text{-0.296(0.321)2 pell lvr on\_1.4kHz=2pell.Standard log scale cubic}_{ij} + \text{0.095(0.019)Pre-fed}_{ij} + \\ & \text{0.019(0.028)2 pell lvr on\_1.Pre-fed}_{ij} + \text{0.118(0.030)4kHz=2pell.Pre-fed}_{ij} + \\ & \text{-0.096(0.041)2 pell lvr on\_1.4kHz=2pell.Pre-fed}_{ij} + \text{-0.209(0.090)UHT}_{j} + \text{0.114(0.023)UHT.2 pell lvr on\_1}_{ij} + e_{ij}\end{aligned}$$

$$\beta_{0j} = \text{-1.158(0.083)} + u_{0j}$$
$$\beta_{1j} = \text{0.500(0.072)} + u_{1j}$$
$$\beta_{2j} = \text{0.415(0.071)} + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} \text{0.036(0.013)} & & \\ \text{0.000(0.004)} & \text{0.006(0.003)} & \\ \text{-0.011(0.009)} & \text{-0.003(0.004)} & \text{0.014(0.010)} \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = \text{0.193(0.003)}$$

*-2*loglikelihood* = 9117.638(7565 of 7565 cases in use)

Equation 7.105 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable.

$$\begin{aligned}\text{RecipRoot\_Rounded\_Latency}_{ij} = {} & \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + \text{-0.472(0.154)Standard log scale cubic}_{ij} + \\ & \text{-0.049(0.029)2 pell lvr on\_1}_{ij} + \text{-1.402(0.088)2 pell lvr on\_1.Standard log scale}_{ij} + \\ & \text{0.055(0.105)2 pell lvr on\_1.Standard log scale quadratic}_{ij} + \text{1.612(0.223)2 pell lvr on\_1.Standard log scale cubic}_{ij} + \\ & \text{0.067(0.101)4kHz=2pell}_{j} + \text{-0.099(0.100)4kHz=2pell.Standard log scale}_{ij} + \\ & \text{-0.281(0.125)4kHz=2pell.Standard log scale quadratic}_{ij} + \text{-0.450(0.229)4kHz=2pell.Standard log scale cubic}_{ij} + \\ & \text{-0.130(0.040)2 pell lvr on\_1.4kHz=2pell}_{ij} + \text{0.443(0.125)2 pell lvr on\_1.4kHz=2pell.Standard log scale}_{ij} + \\ & \text{0.373(0.153)2 pell lvr on\_1.4kHz=2pell.Standard log scale quadratic}_{ij} + \\ & \text{-0.296(0.321)2 pell lvr on\_1.4kHz=2pell.Standard log scale cubic}_{ij} + \text{0.095(0.019)Pre-fed}_{ij} + \\ & \text{0.019(0.028)2 pell lvr on\_1.Pre-fed}_{ij} + \text{0.118(0.030)4kHz=2pell.Pre-fed}_{ij} + \\ & \text{-0.096(0.041)2 pell lvr on\_1.4kHz=2pell.Pre-fed}_{ij} + \text{-0.212(0.090)UHT}_{j} + \text{0.114(0.023)UHT.2 pell lvr on\_1}_{ij} + \\ & \text{-0.041(0.089)Enriched}_{j} + e_{ij}\end{aligned}$$

$$\beta_{0j} = \text{-1.136(0.095)} + u_{0j}$$
$$\beta_{1j} = \text{0.501(0.072)} + u_{1j}$$
$$\beta_{2j} = \text{0.415(0.071)} + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} \text{0.035(0.013)} & & \\ \text{0.000(0.004)} & \text{0.006(0.003)} & \\ \text{-0.010(0.009)} & \text{-0.003(0.004)} & \text{0.014(0.010)} \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = \text{0.193(0.003)}$$

*-2*loglikelihood* = 9117.445(7565 of 7565 cases in use)

Equation 7.106 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable.

$$\begin{aligned}
\text{RecipRoot_Rounded_Latency}_{ij} = {} & \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + \text{-0.472(0.154)Standard log scale cubic}_{ij} + \\
& \text{-0.049(0.029)2 pell lvr on_1}_{ij} + \text{-1.402(0.088)2 pell lvr on_1.Standard log scale}_{ij} + \\
& \text{0.058(0.105)2 pell lvr on_1.Standard log scale quadratic}_{ij} + \text{1.616(0.223)2 pell lvr on_1.Standard log scale cubic}_{ij} + \\
& \text{0.089(0.082)4kHz=2pell}_{j} + \text{-0.098(0.100)4kHz=2pell.Standard log scale}_{ij} + \\
& \text{-0.283(0.126)4kHz=2pell.Standard log scale quadratic}_{ij} + \text{-0.452(0.229)4kHz=2pell.Standard log scale cubic}_{ij} + \\
& \text{-0.131(0.040)2 pell lvr on_1.4kHz=2pell}_{ij} + \text{0.443(0.125)2 pell lvr on_1.4kHz=2pell.Standard log scale}_{ij} + \\
& \text{0.375(0.153)2 pell lvr on_1.4kHz=2pell.Standard log scale quadratic}_{ij} + \\
& \text{-0.300(0.321)2 pell lvr on_1.4kHz=2pell.Standard log scale cubic}_{ij} + \text{0.095(0.019)Pre-fed}_{ij} + \\
& \text{0.019(0.028)2 pell lvr on_1.Pre-fed}_{ij} + \text{0.117(0.030)4kHz=2pell.Pre-fed}_{ij} + \\
& \text{-0.096(0.041)2 pell lvr on_1.4kHz=2pell.Pre-fed}_{ij} + \text{-0.186(0.074)UHT}_{j} + \text{0.114(0.023)UHT.2 pell lvr on_1}_{ij} + \\
& \text{0.225(0.074)Prefed_Session1}_{j} + e_{ij}
\end{aligned}$$

$$\beta_{0j} = \text{-1.282(0.077)} + u_{0j}$$
$$\beta_{1j} = \text{0.500(0.073)} + u_{1j}$$
$$\beta_{2j} = \text{0.415(0.071)} + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ : \ \Omega_u = \begin{bmatrix} \text{0.022(0.008)} & & \\ \text{0.003(0.004)} & \text{0.006(0.003)} & \\ \text{-0.005(0.007)} & \text{-0.003(0.004)} & \text{0.015(0.010)} \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = \text{0.193(0.003)}$$

*-2*loglikelihood* = 9111.855(7565 of 7565 cases in use)

Equation 7.107 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable.

$$\begin{aligned}
\text{RecipRoot_Rounded_Latency}_{ij} = {} & \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + \text{-0.479(0.154)Standard log scale cubic}_{ij} + \\
& \text{-0.044(0.029)2 pell lvr on_1}_{ij} + \text{-1.410(0.088)2 pell lvr on_1.Standard log scale}_{ij} + \\
& \text{0.050(0.105)2 pell lvr on_1.Standard log scale quadratic}_{ij} + \text{1.642(0.223)2 pell lvr on_1.Standard log scale cubic}_{ij} + \\
& \text{0.092(0.082)4kHz=2pell}_{j} + \text{-0.103(0.100)4kHz=2pell.Standard log scale}_{ij} + \\
& \text{-0.280(0.126)4kHz=2pell.Standard log scale quadratic}_{ij} + \text{-0.437(0.229)4kHz=2pell.Standard log scale cubic}_{ij} + \\
& \text{-0.134(0.040)2 pell lvr on_1.4kHz=2pell}_{ij} + \text{0.458(0.125)2 pell lvr on_1.4kHz=2pell.Standard log scale}_{ij} + \\
& \text{0.374(0.153)2 pell lvr on_1.4kHz=2pell.Standard log scale quadratic}_{ij} + \\
& \text{-0.333(0.320)2 pell lvr on_1.4kHz=2pell.Standard log scale cubic}_{ij} + \text{0.095(0.019)Pre-fed}_{ij} + \\
& \text{0.016(0.028)2 pell lvr on_1.Pre-fed}_{ij} + \text{0.116(0.030)4kHz=2pell.Pre-fed}_{ij} + \\
& \text{-0.094(0.041)2 pell lvr on_1.4kHz=2pell.Pre-fed}_{ij} + \text{-0.184(0.074)UHT}_{j} + \text{0.115(0.023)UHT.2 pell lvr on_1}_{ij} + \\
& \text{0.225(0.074)Prefed_Session1}_{j} + \text{0.013(0.003)Session}_{ij} + e_{ij}
\end{aligned}$$

$$\beta_{0j} = \text{-1.285(0.077)} + u_{0j}$$
$$\beta_{1j} = \text{0.501(0.072)} + u_{1j}$$
$$\beta_{2j} = \text{0.417(0.071)} + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ : \ \Omega_u = \begin{bmatrix} \text{0.022(0.008)} & & \\ \text{0.003(0.004)} & \text{0.006(0.003)} & \\ \text{-0.005(0.007)} & \text{-0.003(0.004)} & \text{0.015(0.010)} \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = \text{0.192(0.003)}$$

*-2*loglikelihood* = 9094.042(7565 of 7565 cases in use)

Equation 7.108 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable.

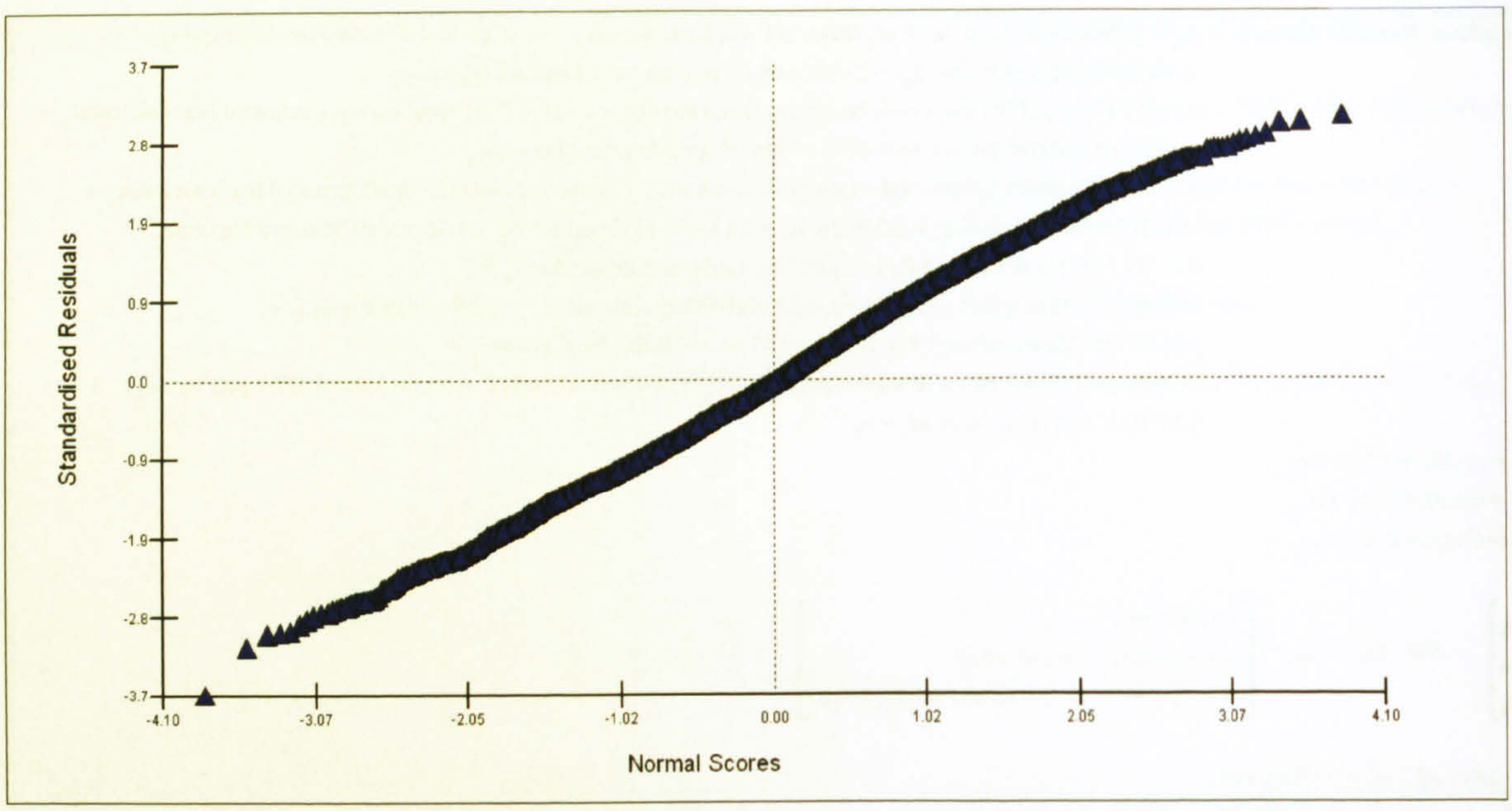

Figure 7.52 The standardised residuals (y-axis) plotted against the Normal scores for the Level 1 (i.e. *trial*) variance, from the model specified in Equation 7.108. NB the conclusions from the model remain the same following the deletion of the outlier at the bottom left of the chart.

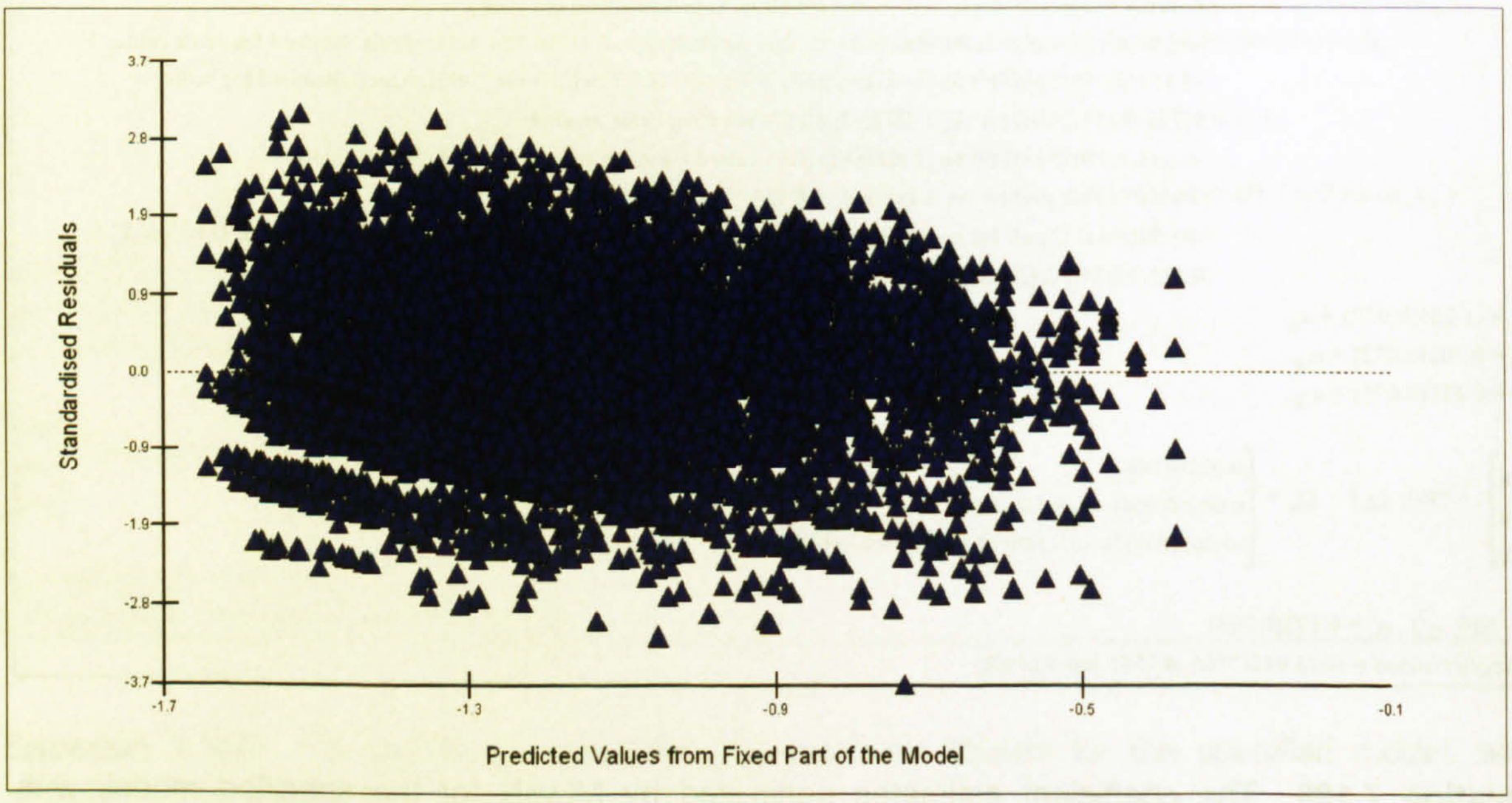

Figure 7.53 The standardised residuals (y-axis) plotted against the predicted values from the fixed part of the model (prior to backtransformation) for the Level 1 (i.e. *trial*) variance, from the model specified in Equation 7.108.

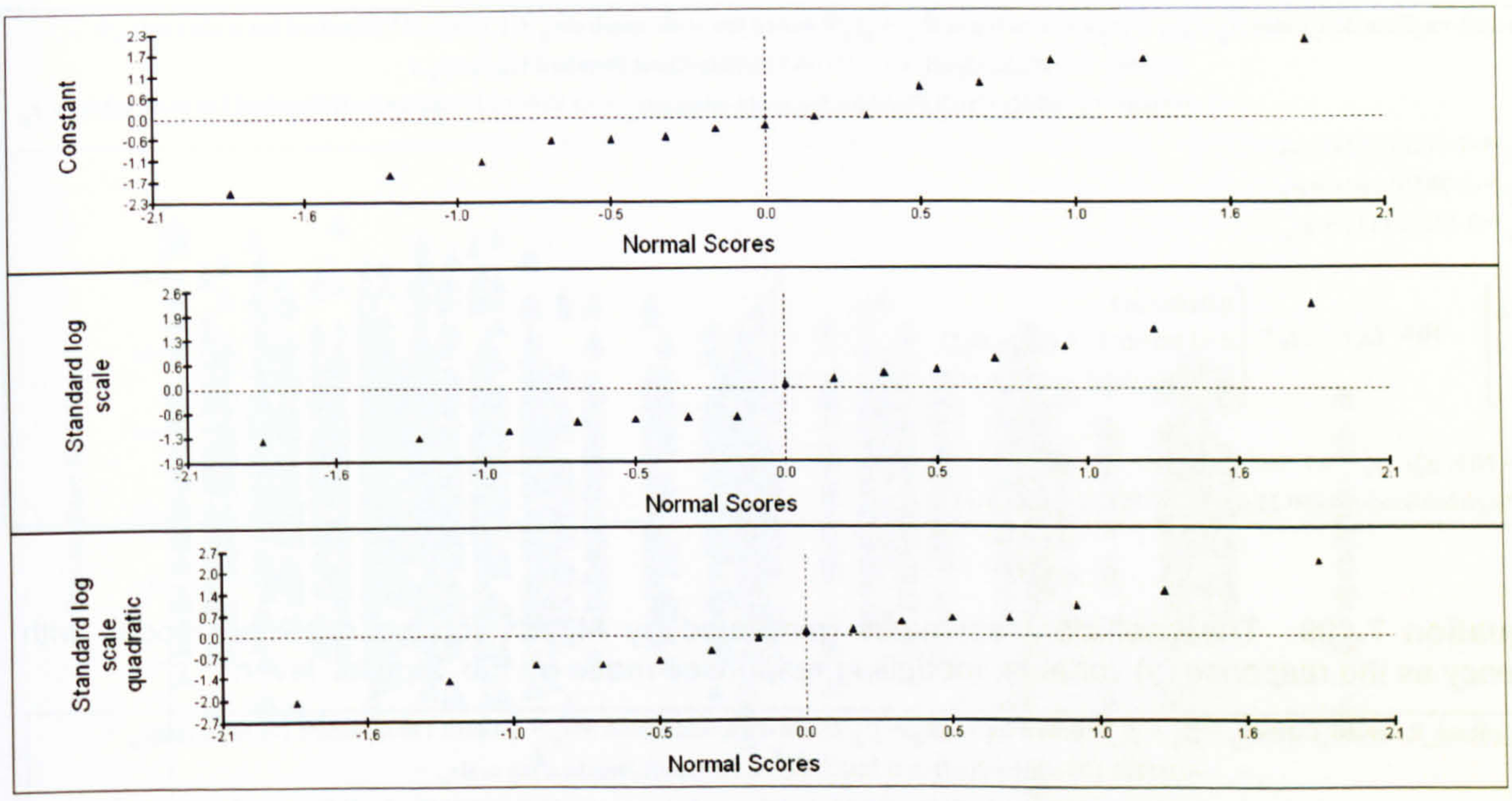

Figure 7.54 The standardised residuals (y-axis) plotted against the Normal scores for the Level 2 (i.e. *subject*) variance, from the model specified in Equation 7.108.

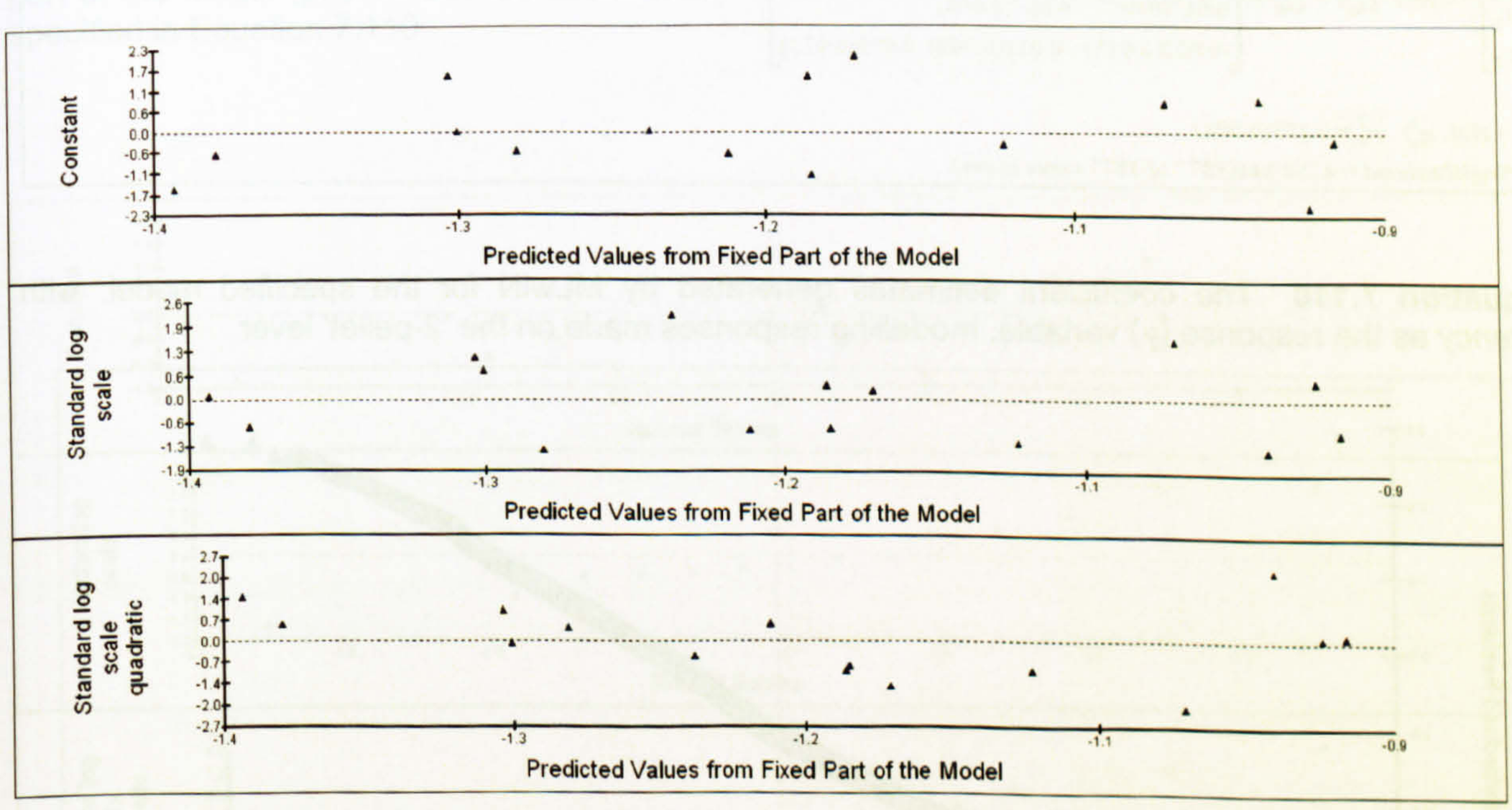

Figure 7.55 The standardised residuals (y-axis) plotted against the predicted values from the fixed part of the model (prior to backtransformation) for the Level 2 (i.e. *subject*) variance, from the model specified in Equation 7.108.

$$\text{RecipRoot\_Rounded\_Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + 1.109(0.155)\text{Standard log scale cubic}_{ij} + -0.039(0.106)\text{4kHz=2pell}_{j} + 0.335(0.093)\text{4kHz=2pell.Standard log scale}_{ij} + 0.064(0.147)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + -0.686(0.215)\text{4kHz=2pell.Standard log scale cubic}_{ij} + e_{ij}$$

$$\beta_{0j} = -1.212(0.073) + u_{0j}$$
$$\beta_{1j} = -0.885(0.063) + u_{1j}$$
$$\beta_{2j} = 0.491(0.111) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.040(0.015) & & \\ 0.014(0.007) & 0.008(0.005) & \\ -0.024(0.015) & -0.013(0.009) & 0.040(0.023) \end{bmatrix}$$

$$e_{ij} \sim N(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.176(0.004)$$

*-2*loglikelihood* = 4350.213(3877 of 3877 cases in use)

Equation 7.109 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable, modelling responses made on the '2-pellet' lever.

$$\text{RecipRoot\_Rounded\_Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + 1.123(0.153)\text{Standard log scale cubic}_{ij} + -0.041(0.107)\text{4kHz=2pell}_{j} + 0.340(0.090)\text{4kHz=2pell.Standard log scale}_{ij} + 0.074(0.145)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + -0.722(0.212)\text{4kHz=2pell.Standard log scale cubic}_{ij} + 0.122(0.013)\text{Pre-fed}_{ij} + e_{ij}$$

$$\beta_{0j} = -1.268(0.074) + u_{0j}$$
$$\beta_{1j} = -0.879(0.062) + u_{1j}$$
$$\beta_{2j} = 0.476(0.110) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim N(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.041(0.016) & & \\ 0.013(0.007) & 0.007(0.004) & \\ -0.022(0.015) & -0.011(0.008) & 0.039(0.023) \end{bmatrix}$$

$$e_{ij} \sim N(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.173(0.004)$$

*-2*loglikelihood* = 4268.844(3877 of 3877 cases in use)

Equation 7.110 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable, modelling responses made on the '2-pellet' lever.

Figure 7.56 The standardised residuals (y-axis) plotted against the Normal scores for the Level 1 (i.e. *trial*) variance, from the model specified in Equation 7.110.

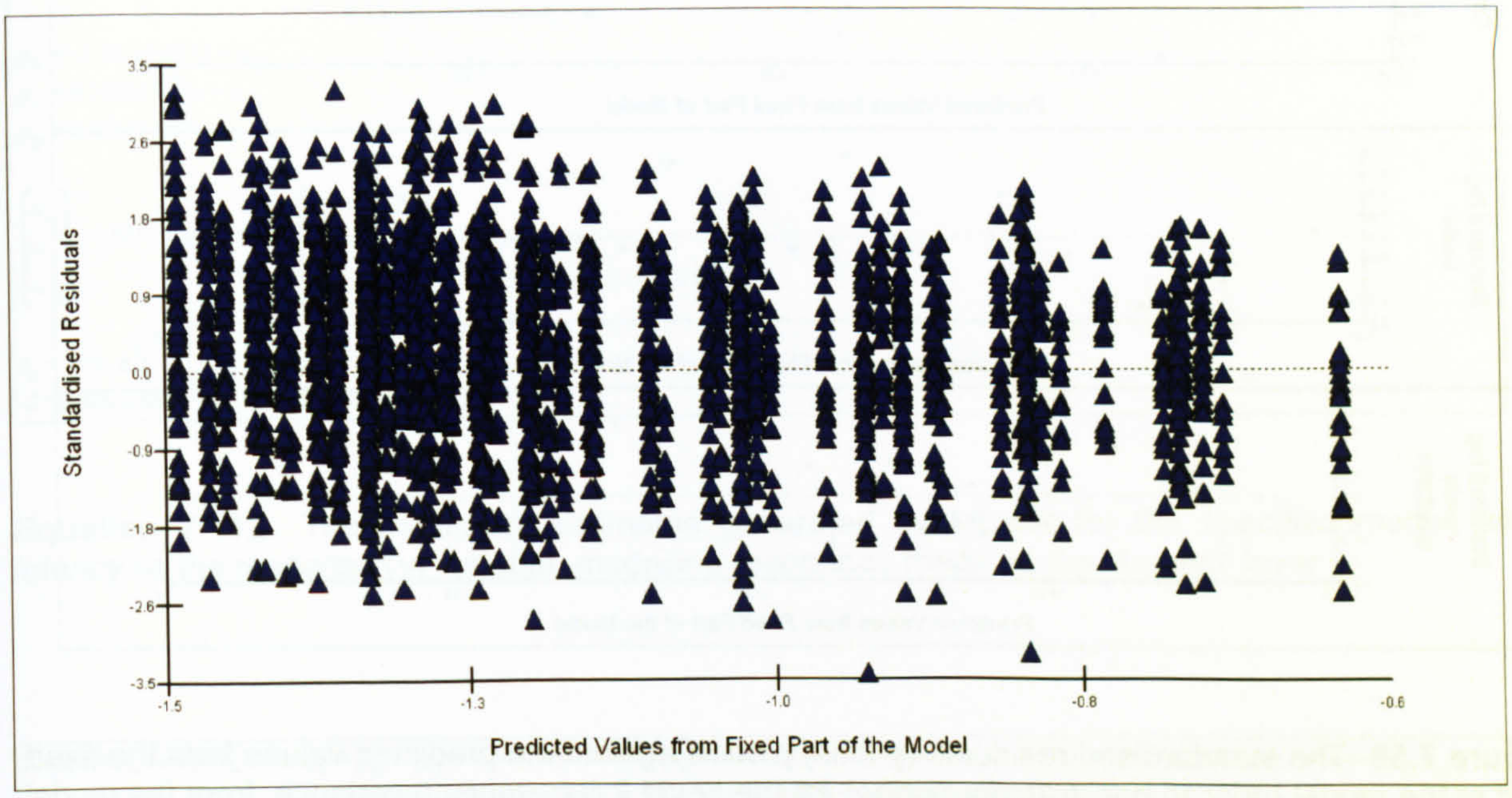

Figure 7.57 The standardised residuals (y-axis) plotted against the predicted values from the fixed part of the model (prior to backtransformation) for the Level 1 (i.e. *trial*) variance, from the model specified in Equation 7.110.

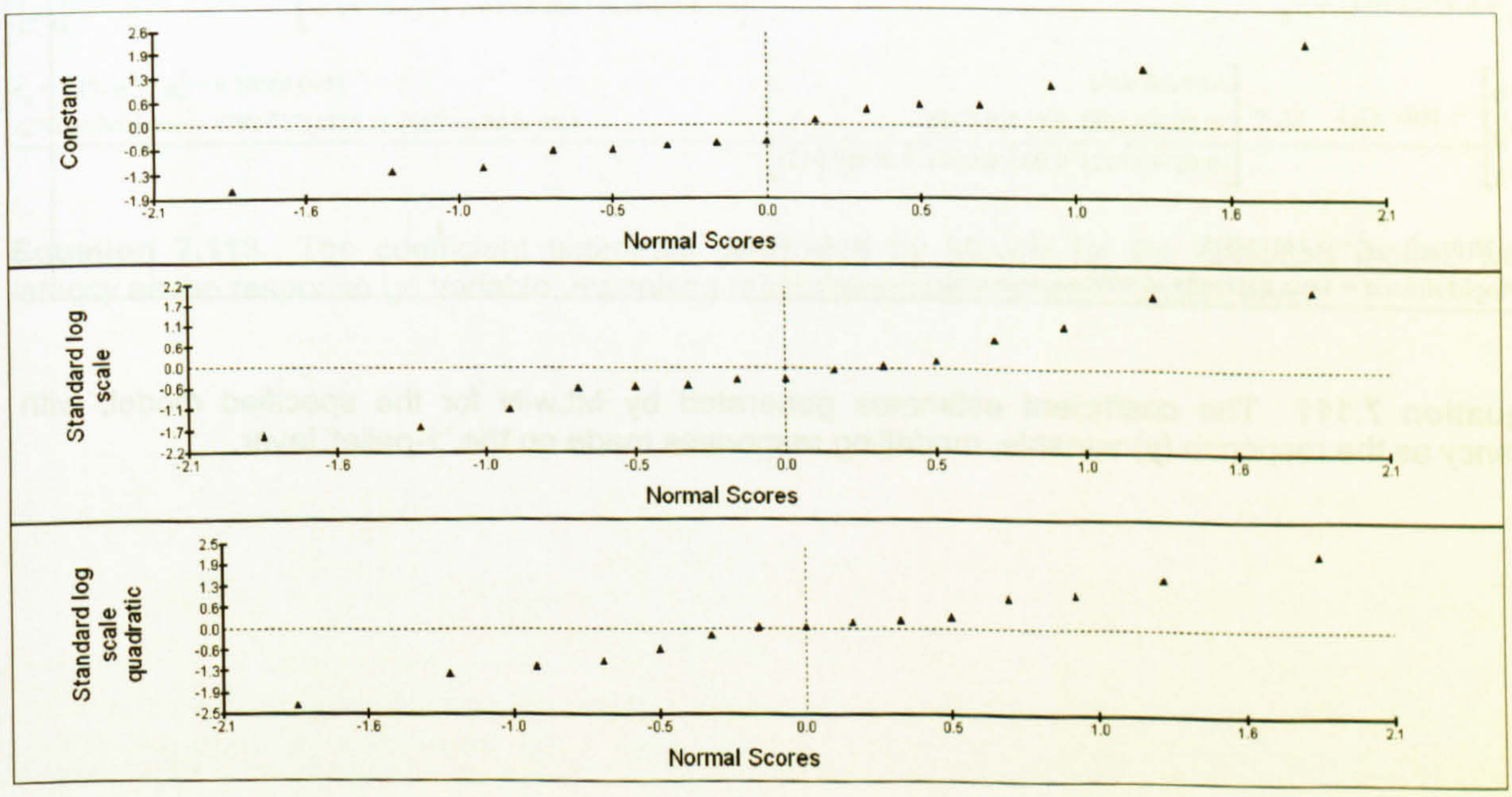

Figure 7.58 The standardised residuals (y-axis) plotted against the Normal scores for the Level 2 (i.e. *subject*) variance, from the model specified in Equation 7.110.

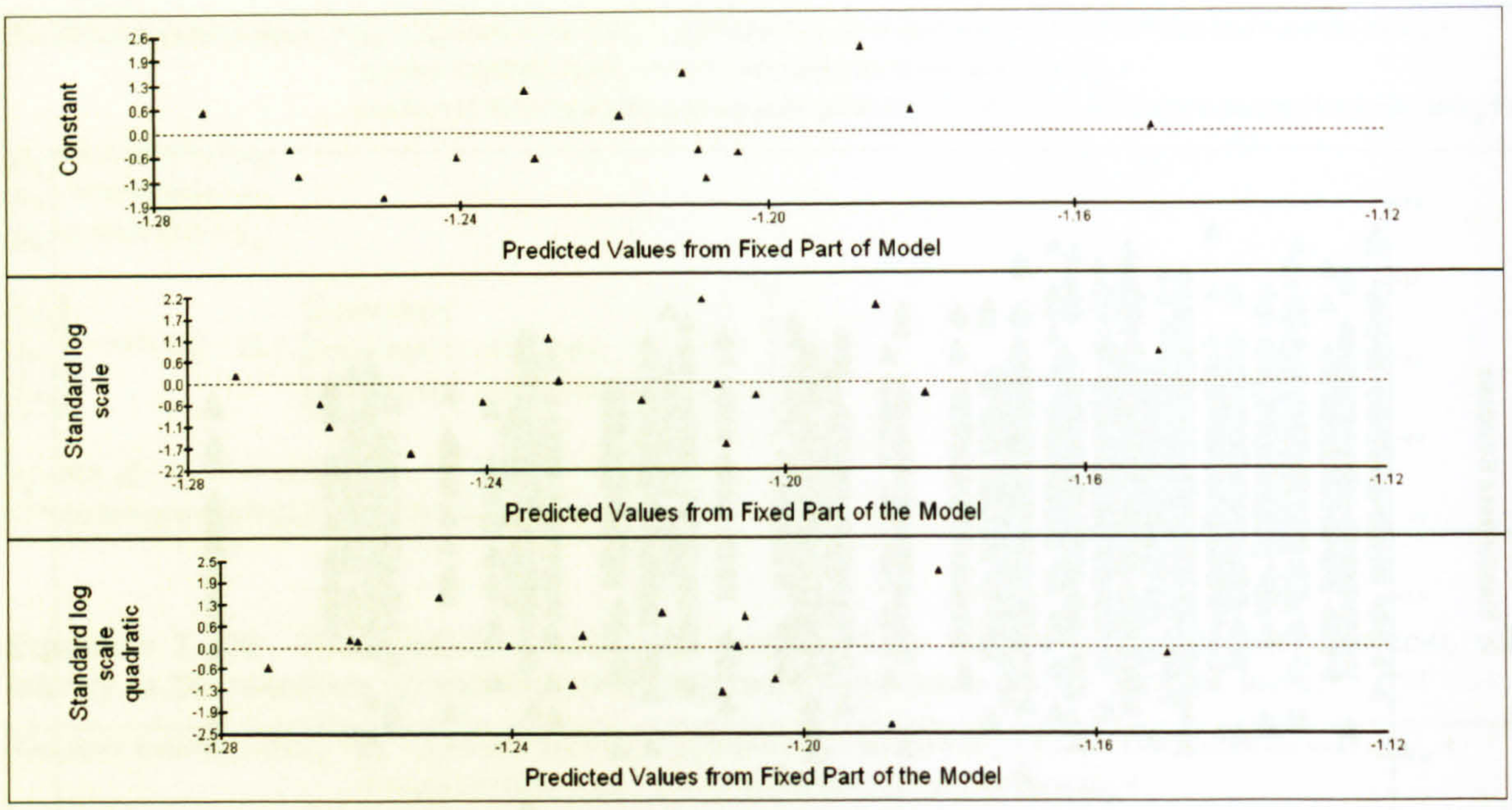

Figure 7.59 The standardised residuals (y-axis) plotted against the predicted values from the fixed part of the model (prior to backtransformation) for the Level 2 (i.e. *subject*) variance, from the model specified in Equation 7.110.

$$\begin{aligned}
\text{RecipRoot_Rounded_Latency}_{ij} = {} & \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + \\
& -0.534(0.161)\text{Standard log scale cubic}_{ij} + 0.139(0.121)\text{4kHz=2pell}_{j} + \\
& -0.102(0.110)\text{4kHz=2pell.Standard log scale}_{ij} + -0.259(0.119)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + \\
& -0.378(0.241)\text{4kHz=2pell.Standard log scale cubic}_{ij} + e_{ij}
\end{aligned}$$

$$\beta_{0j} = -1.213(0.082) + u_{0j}$$
$$\beta_{1j} = 0.535(0.080) + u_{1j}$$
$$\beta_{2j} = 0.428(0.064) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.052(0.020) & & \\ -0.007(0.008) & 0.011(0.006) & \\ -0.017(0.012) & 0.005(0.006) & 0.004(0.011) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.211(0.005)$$

$-2 * loglikelihood = 4802.801(3688 \text{ of } 3688 \text{ cases in use})$

Equation 7.111 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable, modelling responses made on the '1-pellet' lever.

$$\text{RecipRoot\_Rounded\_Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -0.503(0.160)\text{Standard log scale cubic}_{ij} + 0.137(0.122)\text{4kHz=2pell}_{j} + -0.101(0.108)\text{4kHz=2pell.Standard log scale}_{ij} + -0.246(0.118)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + -0.377(0.238)\text{4kHz=2pell.Standard log scale cubic}_{ij} + 0.136(0.015)\text{Pre-fed}_{ij} + e_{ij}$$

$$\beta_{0j} = -1.281(0.083) + u_{0j}$$
$$\beta_{1j} = 0.515(0.078) + u_{1j}$$
$$\beta_{2j} = 0.425(0.063) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.053(0.020) & & \\ -0.006(0.008) & 0.009(0.005) & \\ -0.017(0.012) & 0.005(0.006) & 0.004(0.011) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.207(0.005)$$

*-2*loglikelihood* = 4722.406(3688 of 3688 cases in use)

Equation 7.112 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable, modelling responses made on the '1-pellet' lever.

$$\text{RecipRoot\_Rounded\_Latency}_{ij} = \beta_{0j} + \beta_{1j}\text{Standard log scale}_{ij} + \beta_{2j}\text{Standard log scale quadratic}_{ij} + -0.513(0.159)\text{Standard log scale cubic}_{ij} + 0.079(0.122)\text{4kHz=2pell}_{j} + -0.117(0.107)\text{4kHz=2pell.Standard log scale}_{ij} + -0.243(0.118)\text{4kHz=2pell.Standard log scale quadratic}_{ij} + -0.351(0.238)\text{4kHz=2pell.Standard log scale cubic}_{ij} + 0.092(0.019)\text{Pre-fed}_{ij} + 0.114(0.031)\text{4kHz=2pell.Pre-fed\#1}_{ij} + e_{ij}$$

$$\beta_{0j} = -1.258(0.083) + u_{0j}$$
$$\beta_{1j} = 0.522(0.077) + u_{1j}$$
$$\beta_{2j} = 0.426(0.063) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \\ u_{2j} \end{bmatrix} \sim \text{N}(0,\ \Omega_u) \ :\ \Omega_u = \begin{bmatrix} 0.052(0.020) & & \\ -0.007(0.007) & 0.009(0.005) & \\ -0.017(0.012) & 0.005(0.006) & 0.004(0.011) \end{bmatrix}$$

$$e_{ij} \sim \text{N}(0,\ \sigma_e^2) \quad \sigma_e^2 = 0.206(0.005)$$

*-2*loglikelihood* = 4708.717(3688 of 3688 cases in use)

Equation 7.113 The coefficient estimates generated by MLwiN for the specified model, with *latency* as the response (*y*) variable, modelling responses made on the '1-pellet' lever.

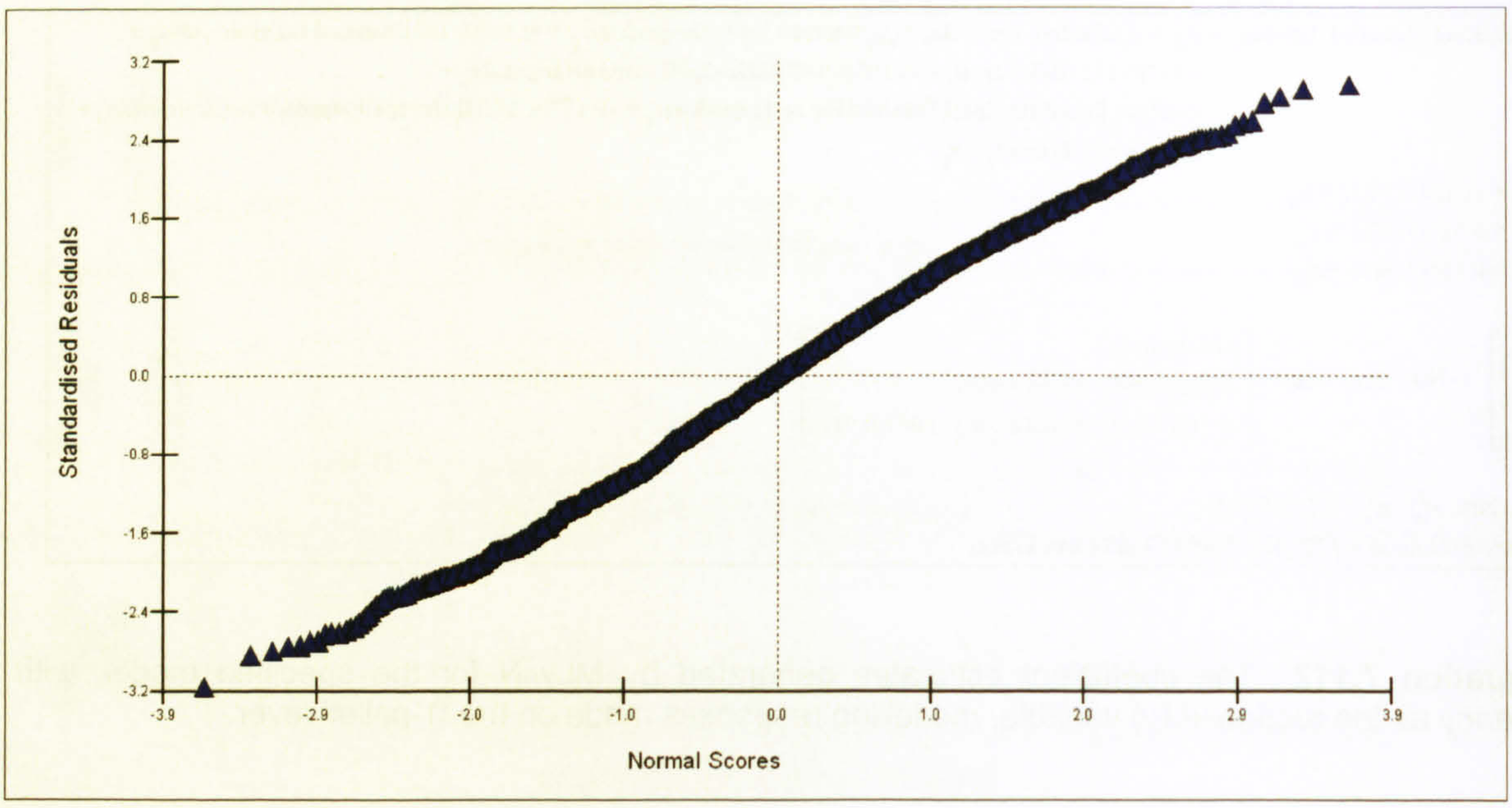

Figure 7.60 The standardised residuals (y-axis) plotted against the Normal scores for the Level 1 (i.e. *trial*) variance, from the model specified in Equation 7.113.

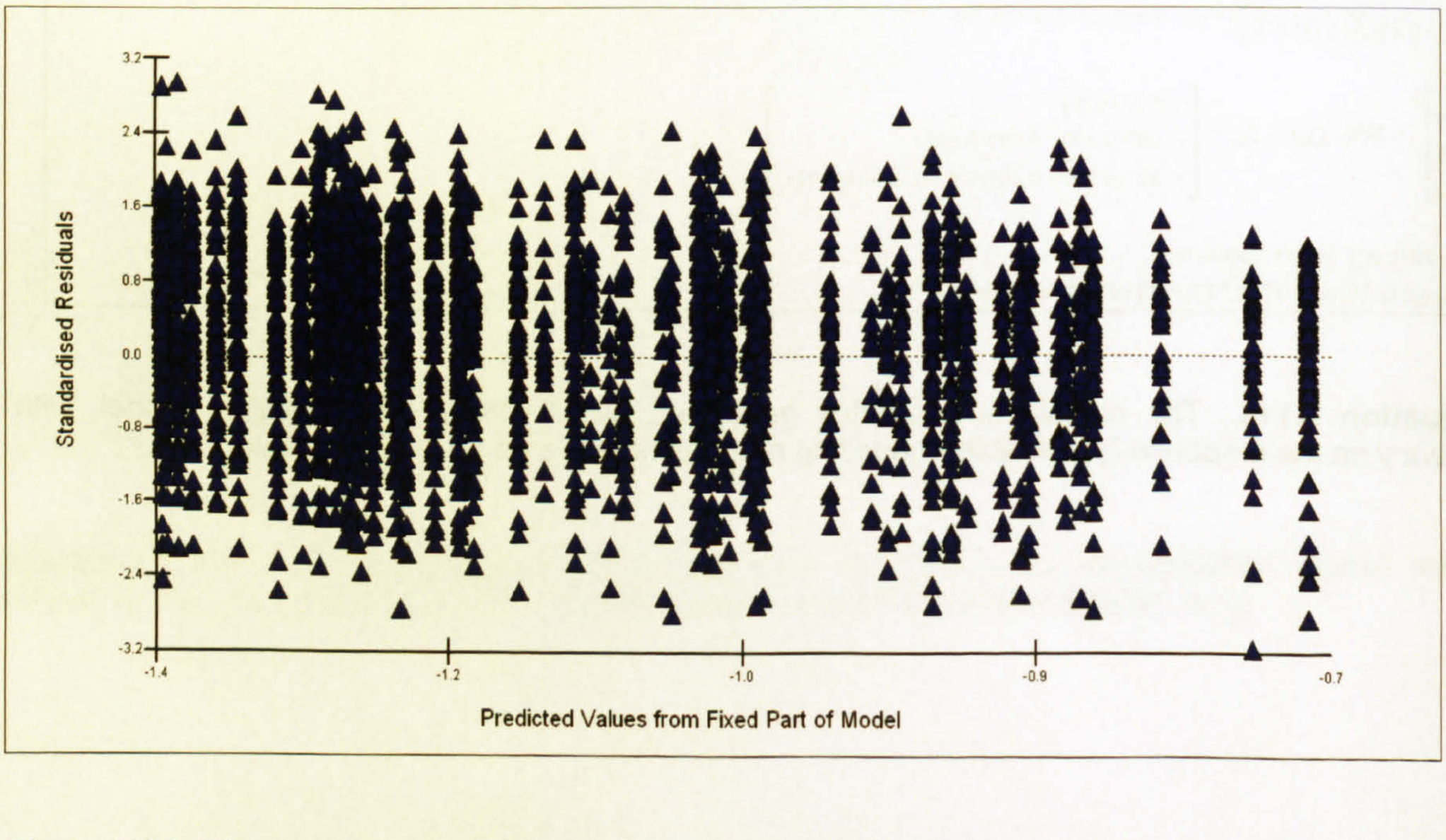

Figure 7.61 The standardised residuals (y-axis) plotted against the predicted values from the fixed part of the model (prior to backtransformation) for the Level 1 (i.e. *trial*) variance, from the model specified in Equation 7.113.

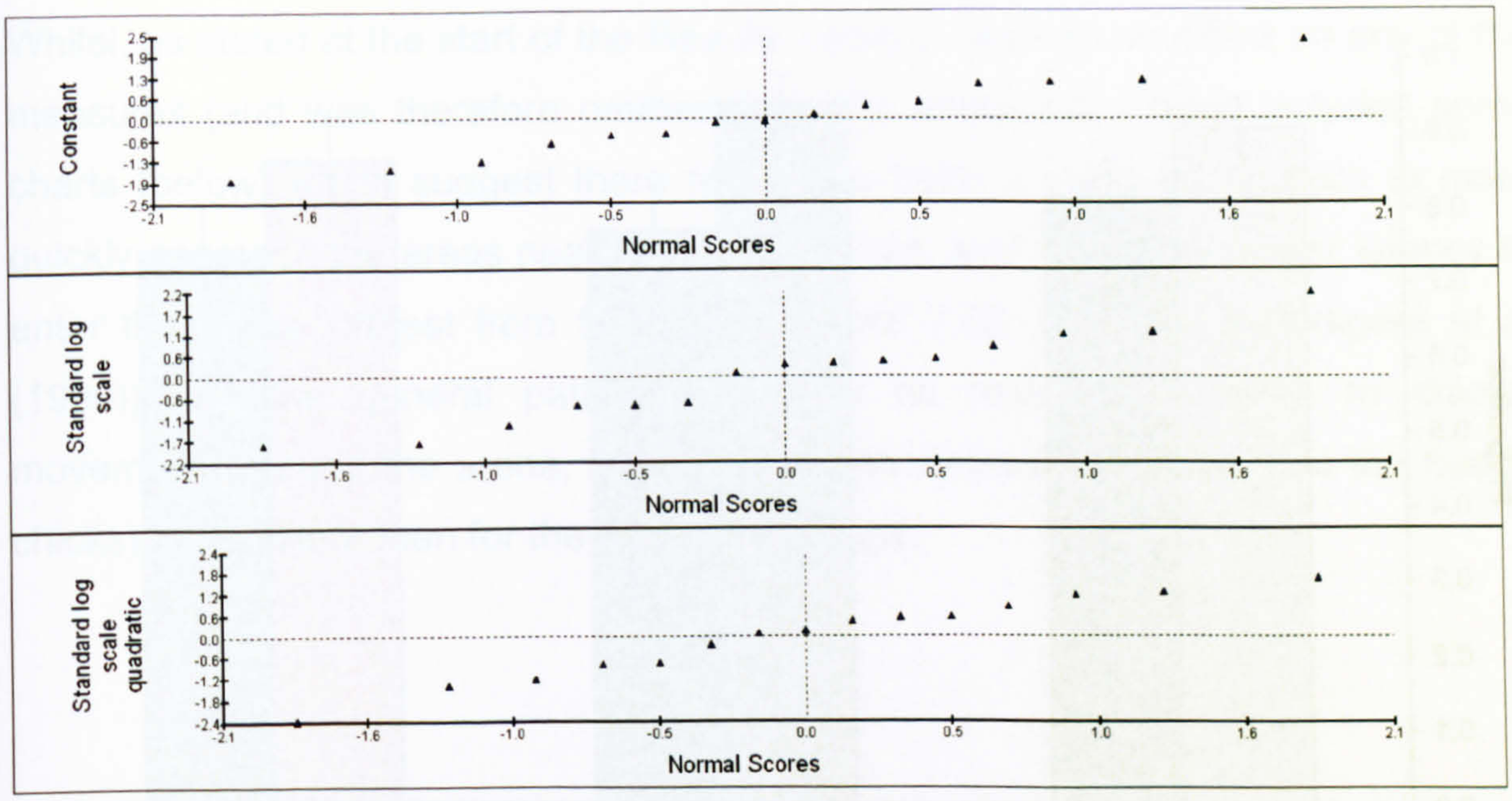

Figure 7.62 The standardised residuals (y-axis) plotted against the Normal scores for the Level 2 (i.e. *subject*) variance, from the model specified in Equation 7.113.

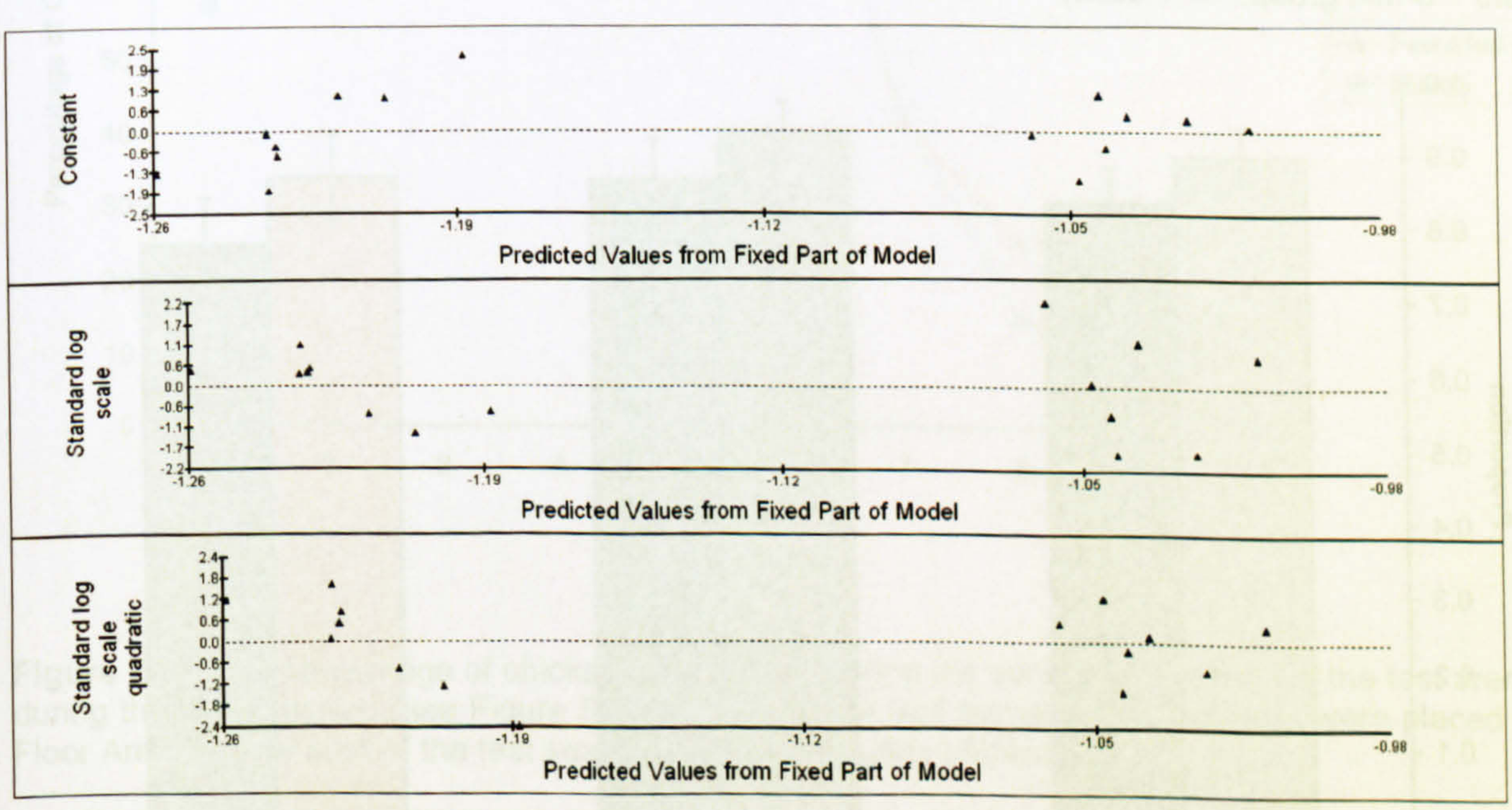

Figure 7.63 The standardised residuals (y-axis) plotted against the predicted values from the fixed part of the model (prior to backtransformation) for the Level 2 (i.e. *subject*) variance, from the model specified in Equation 7.113.

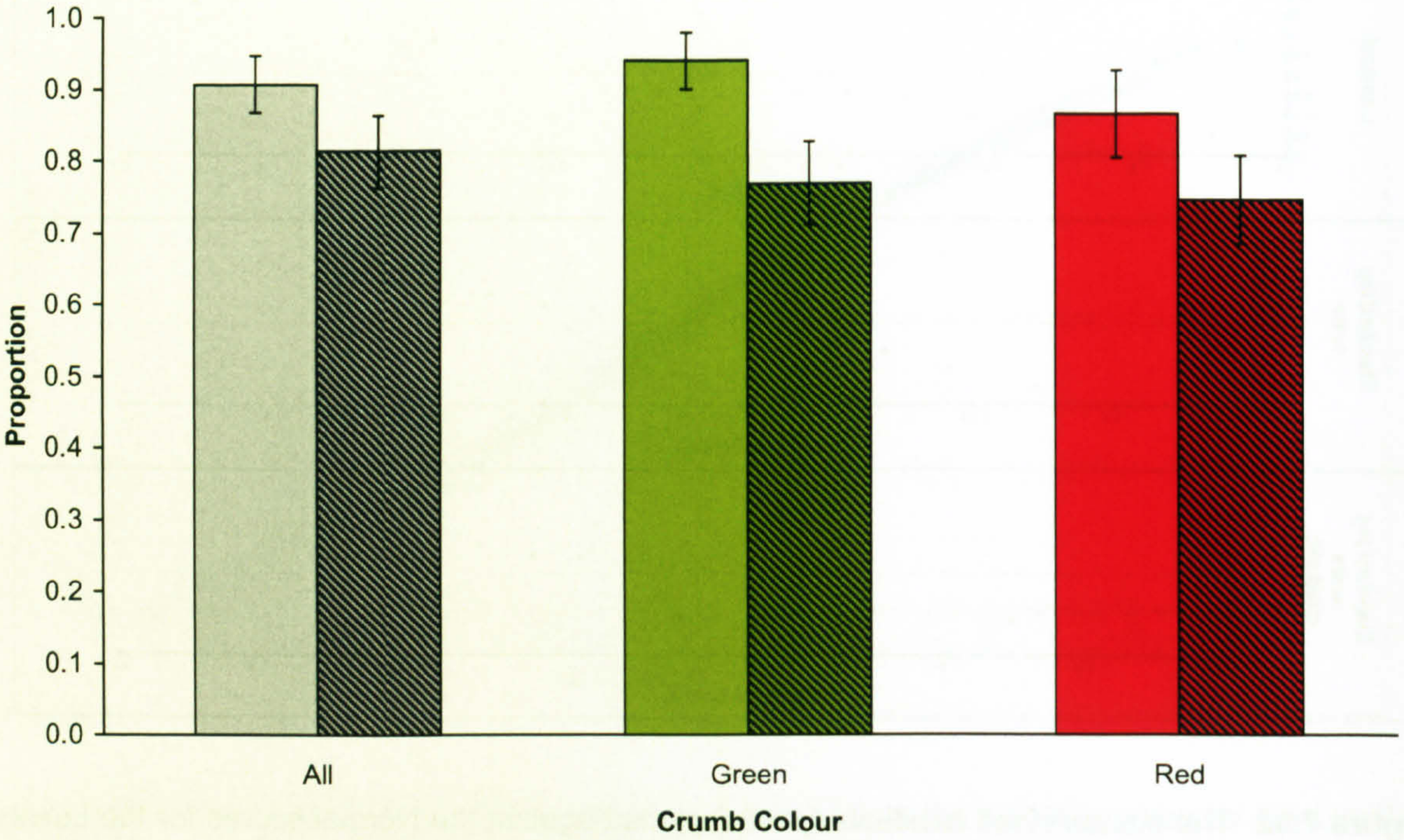

Figure 7.64 The mean proportion of crumb attacks in which the crumb was eaten (rather than being pecked & rejected) for the first eight attacks only (non-shaded bars = 0-min group, shaded bars = 3-min group; +/- 1 SEM).

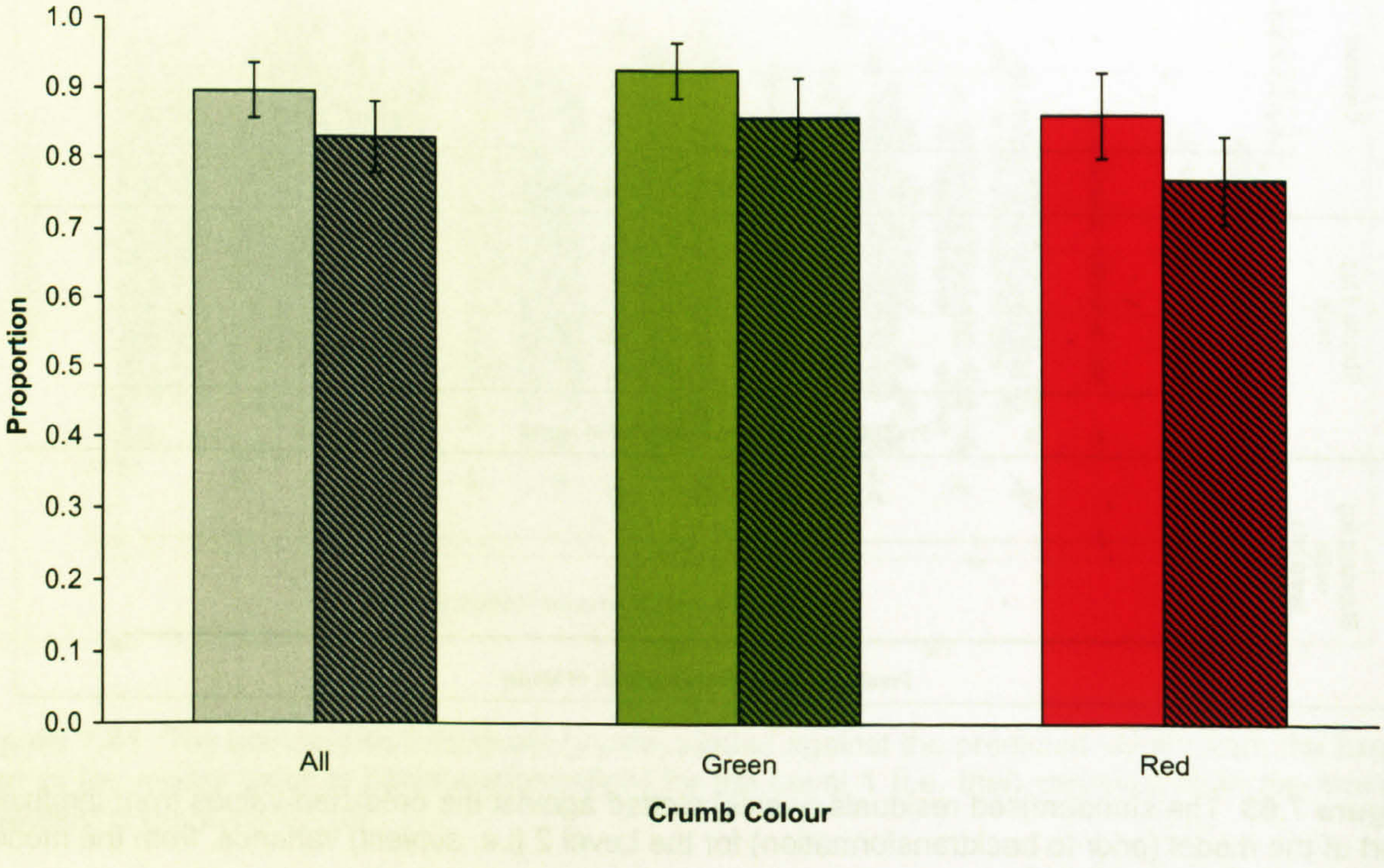

Figure 7.65 The mean proportion of crumb attacks in which the crumb was eaten (rather than being pecked & rejected) for the last eight attacks only (non-shaded bars = 0-min group, shaded bars = 3-min group; +/- 1 SEM).

Whilst, as stated at the start of the Results section, sex had no effect on any of the measures (and was therefore omitted from the analyses), I have included some charts (below) which suggest there may have been a trend for females to more quickly approach the areas near the buddy chicks, and to have a slower latency to enter the areas furthest from them (see Figure 7.68; also see Vallortigara et al (1990)), but the general pattern (i.e. from all measures relating to chicks' movements around the arena, in particular with respect to proximity to the buddy chicks) is less clear than for the treatment groups.

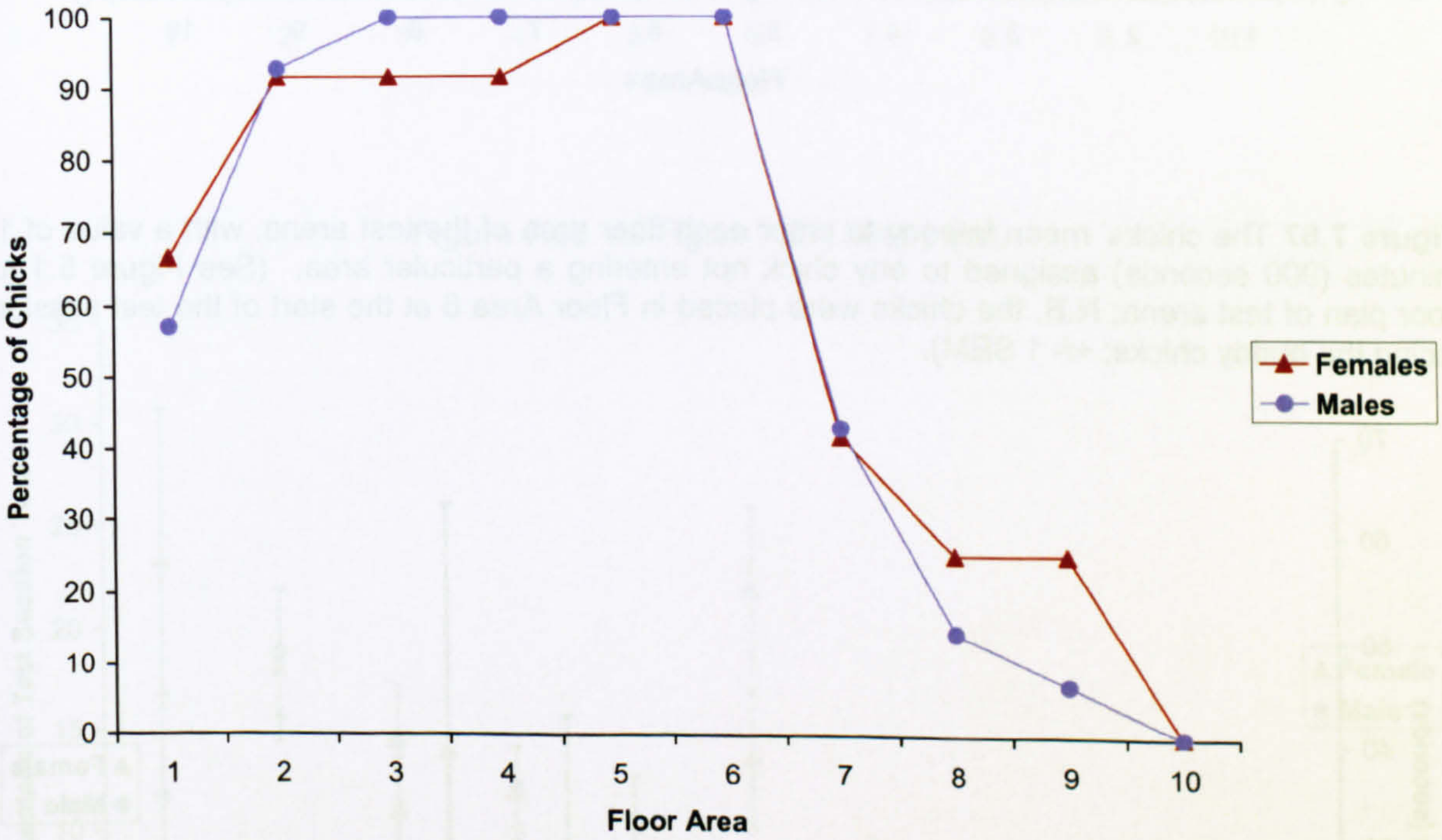

Figure 7.66 The percentage of chicks of each sex entering the various floor areas of the test arena during their test session (see Figure 5.1 for floor plan of test arena; N.B. the chicks were placed in Floor Area 6 at the start of the test session, facing the buddy chicks).

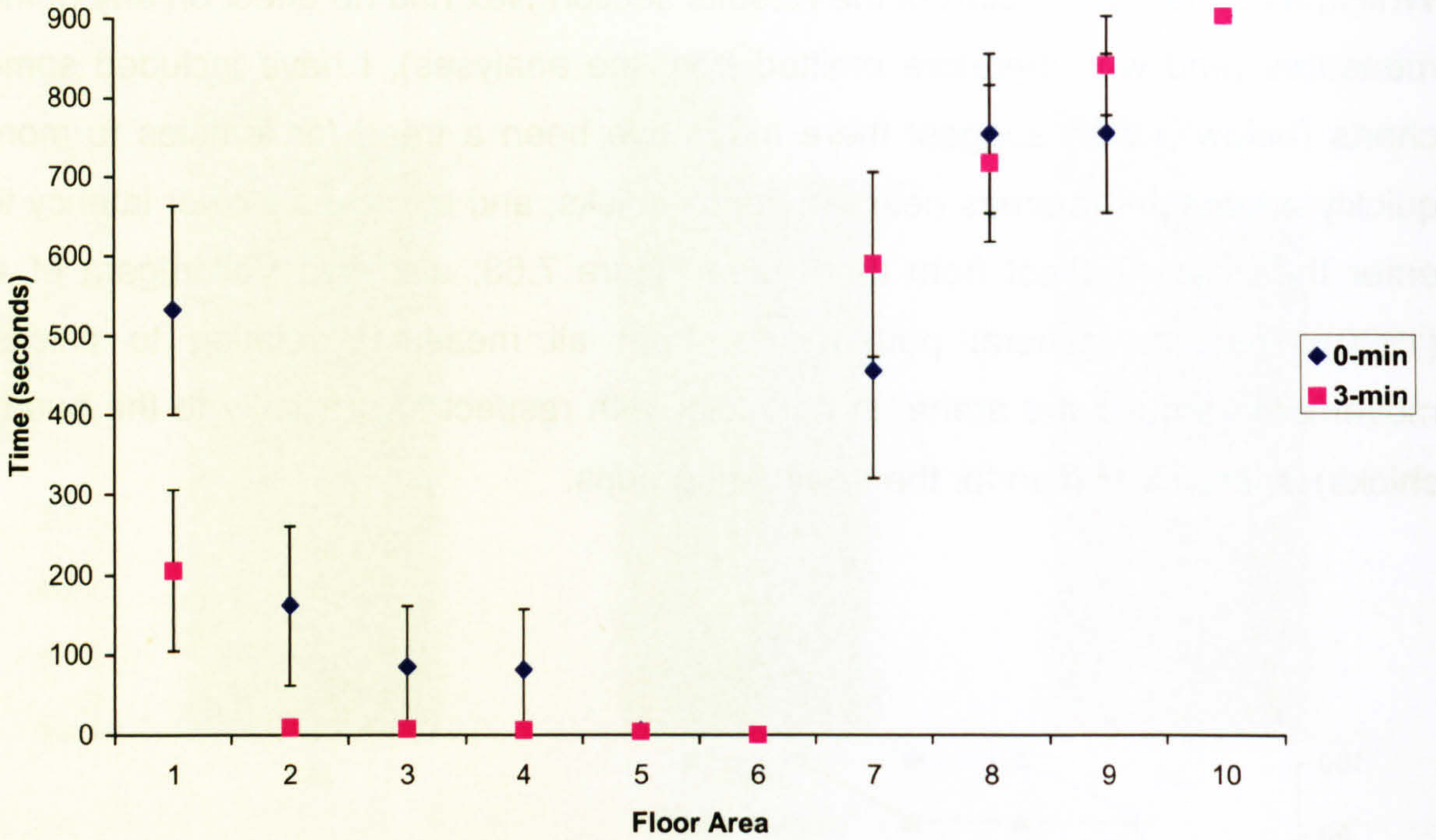

Figure 7.67 The chicks' mean latency to enter each floor area of the test arena, with a value of 15 minutes (900 seconds) assigned to any chick not entering a particular area. (See Figure 5.1 for floor plan of test arena; N.B. the chicks were placed in Floor Area 6 at the start of the test session, facing the buddy chicks; +/- 1 SEM).

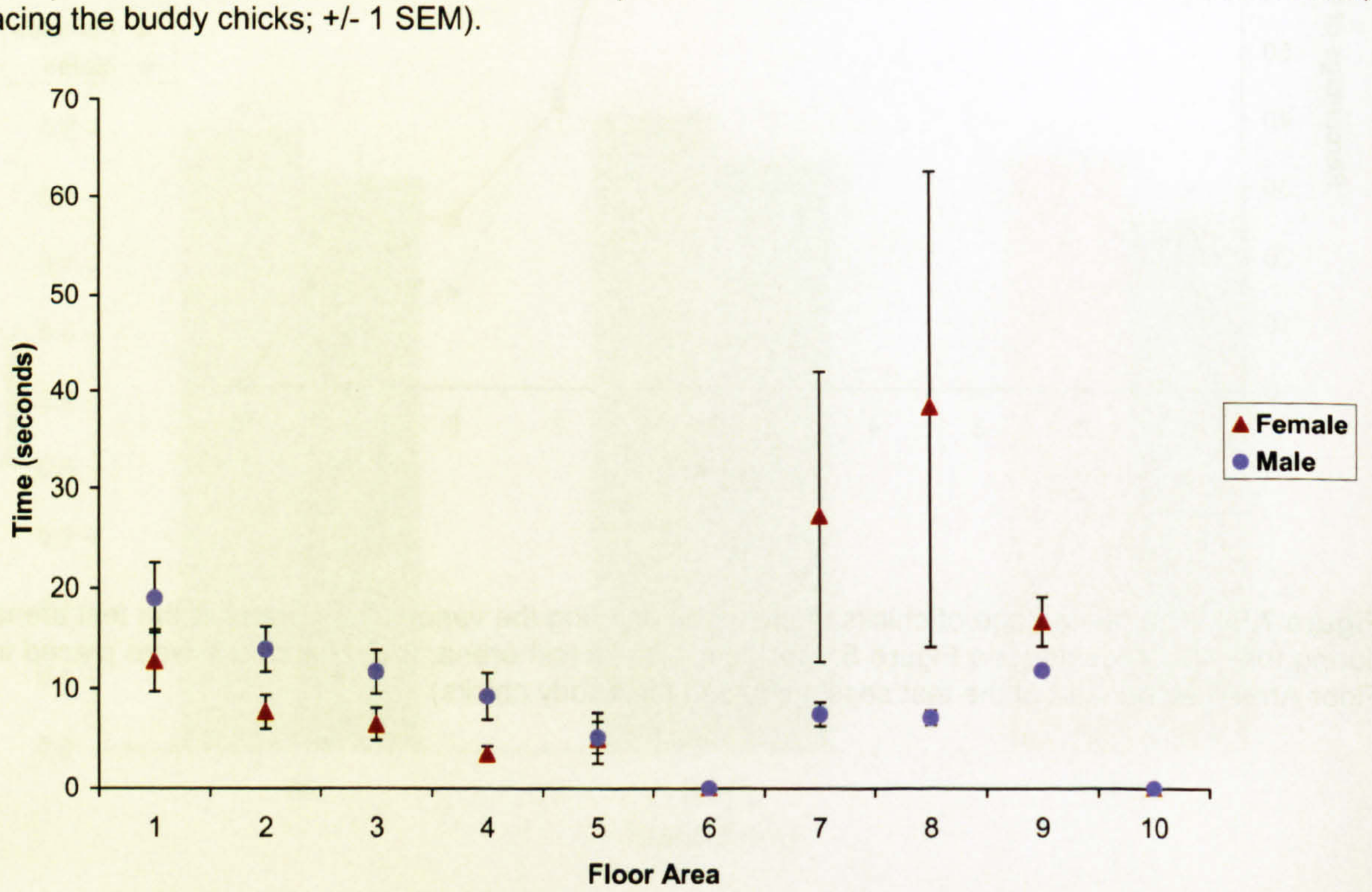

Figure 7.68 The chicks' mean latency to enter each floor area of the test arena, by sex (see Figure 5.1 for floor plan of test arena; N.B. the chicks were placed in Floor Area 6 at the start of the test session, facing the buddy chicks; data is taken only from chicks entering a particular floor area; +/- 1 SEM).

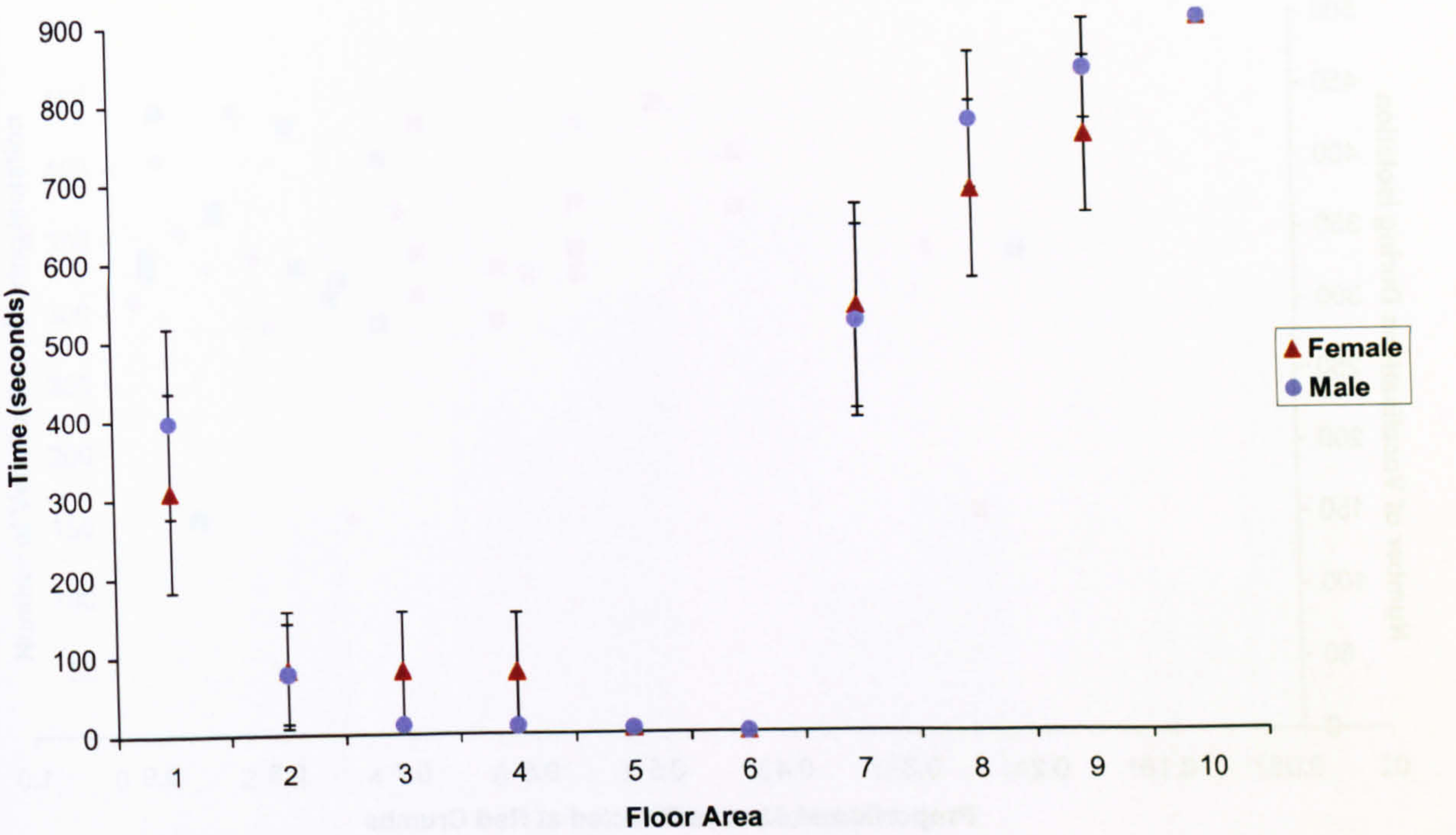

Figure 7.69 As Figure 7.67, but by sex.

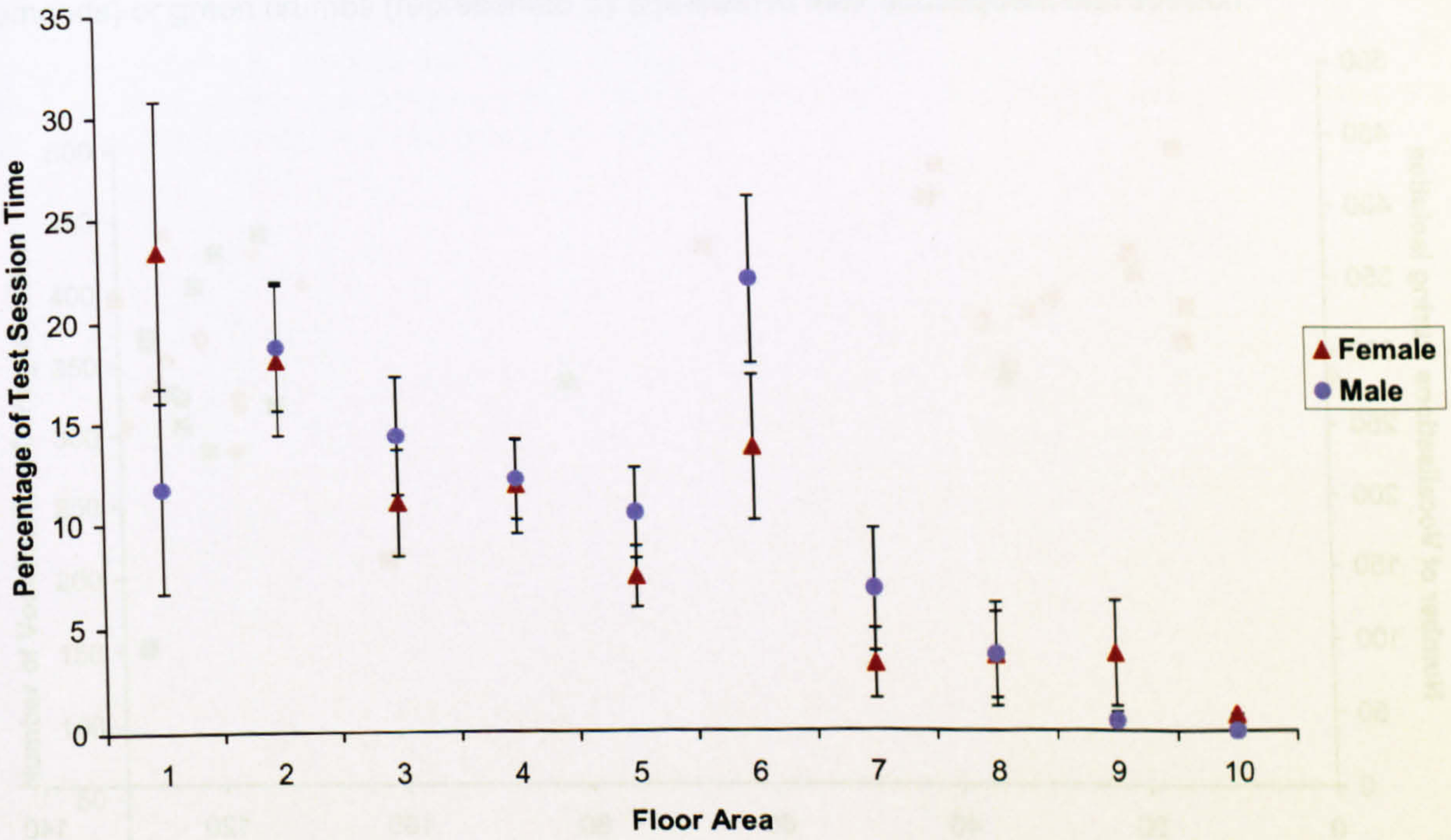

Figure 7.70 The mean percentage of test session time the chicks spent in each floor area of the test arena, by sex (see Figure 5.1 for floor plan of test arena; N.B. the chicks were placed in Floor Area 6 at the start of the test session, facing the buddy chicks; test session duration differed between chicks, and was determined by their latency to attack 16 crumbs; +/- 1 SEM).

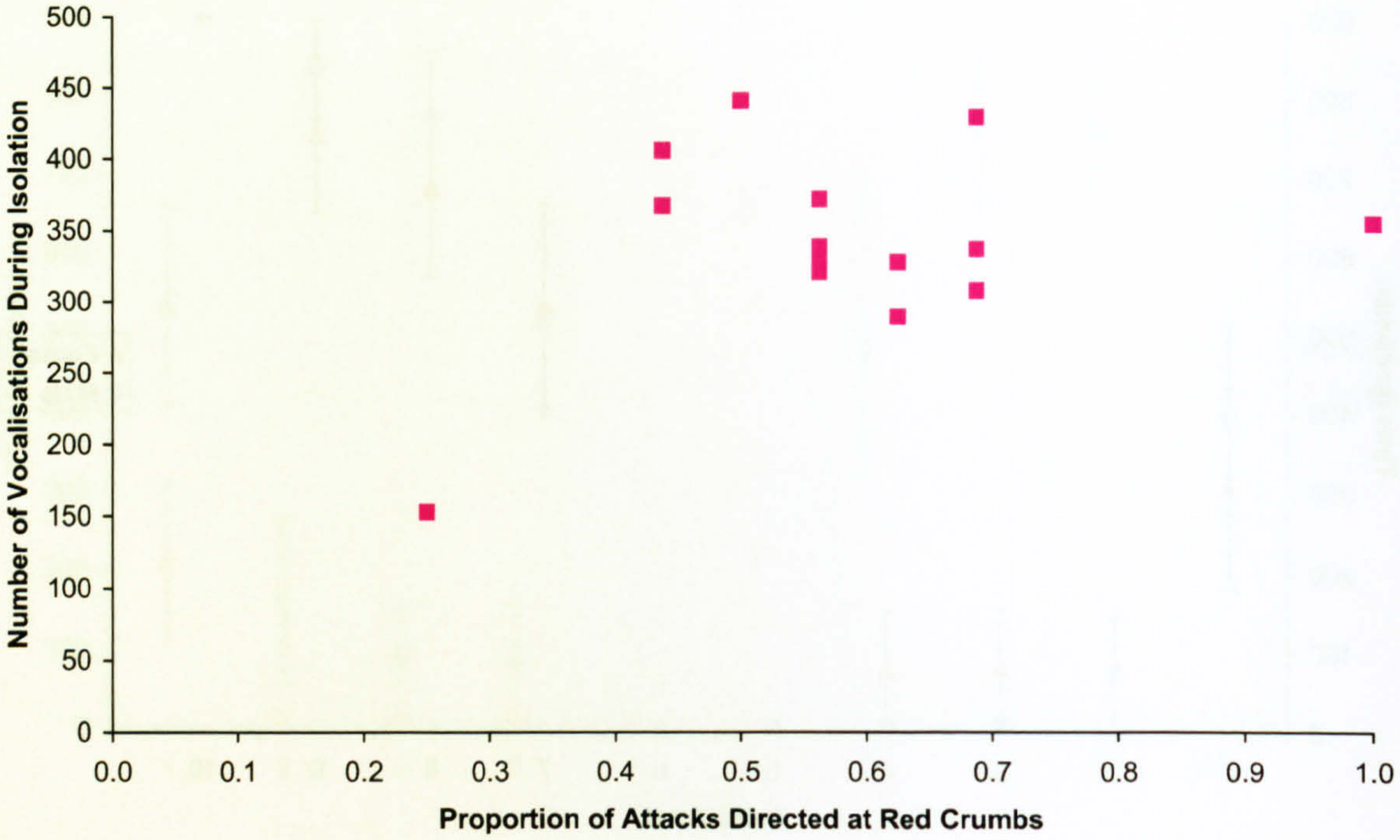

Figure 7.71 The number of vocalisations recorded from the chicks in the 3-min group during their three minute period of social isolation, against the proportion of attacks which were directed towards red crumbs in their subsequent test session.

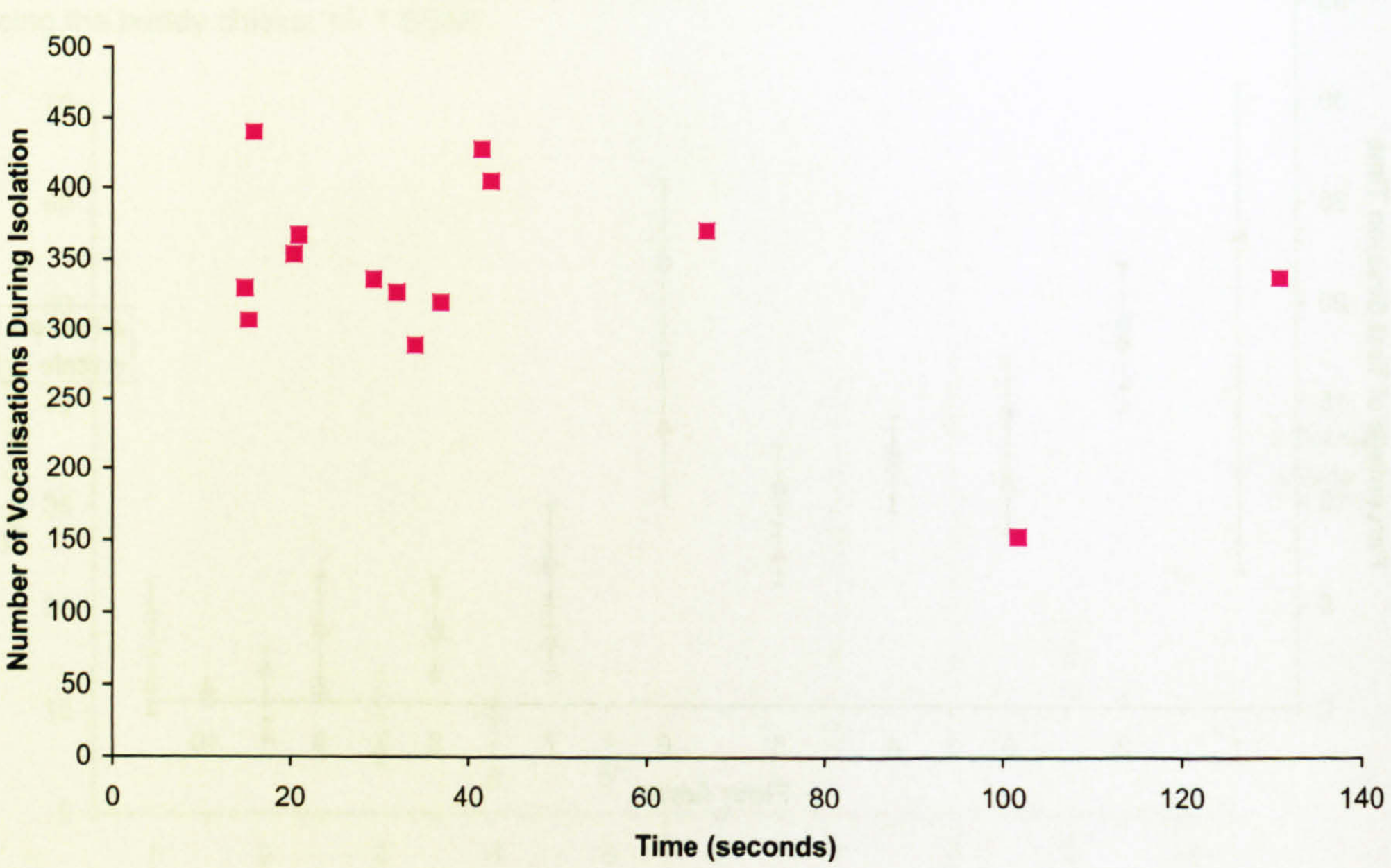

Figure 7.72 The number of vocalisations recorded from the chicks in the 3-min group during their three minute period of social isolation, against the latency to attack 16 crumbs (the criterion for test session termination) in their subsequent test session.

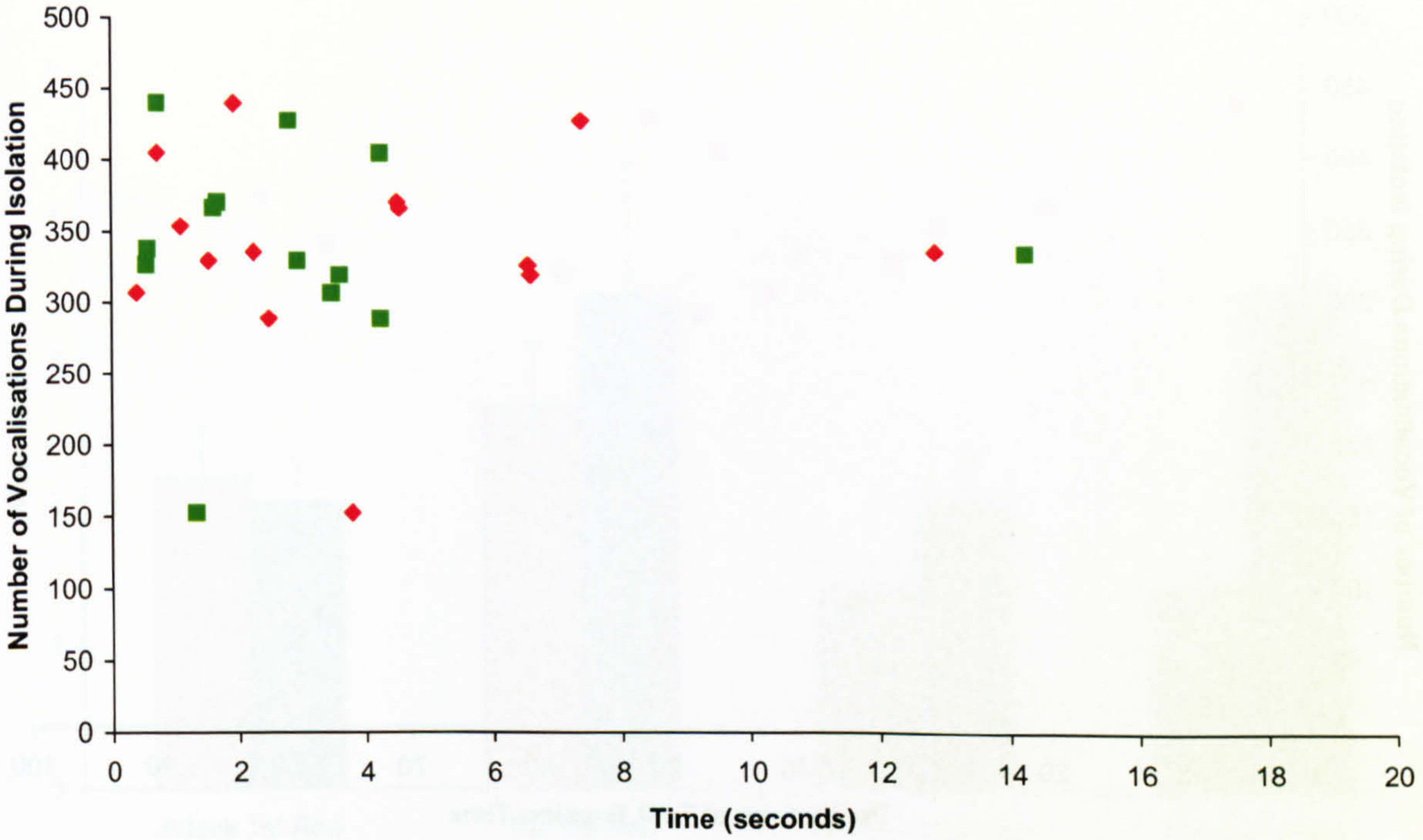

Figure 7.73 The number of vocalisations recorded from the chicks in the 3-min group during their three minute period of social isolation, against their latency to attack the first red (represented by diamonds) or green crumbs (represented by squares) in their subsequent test session.

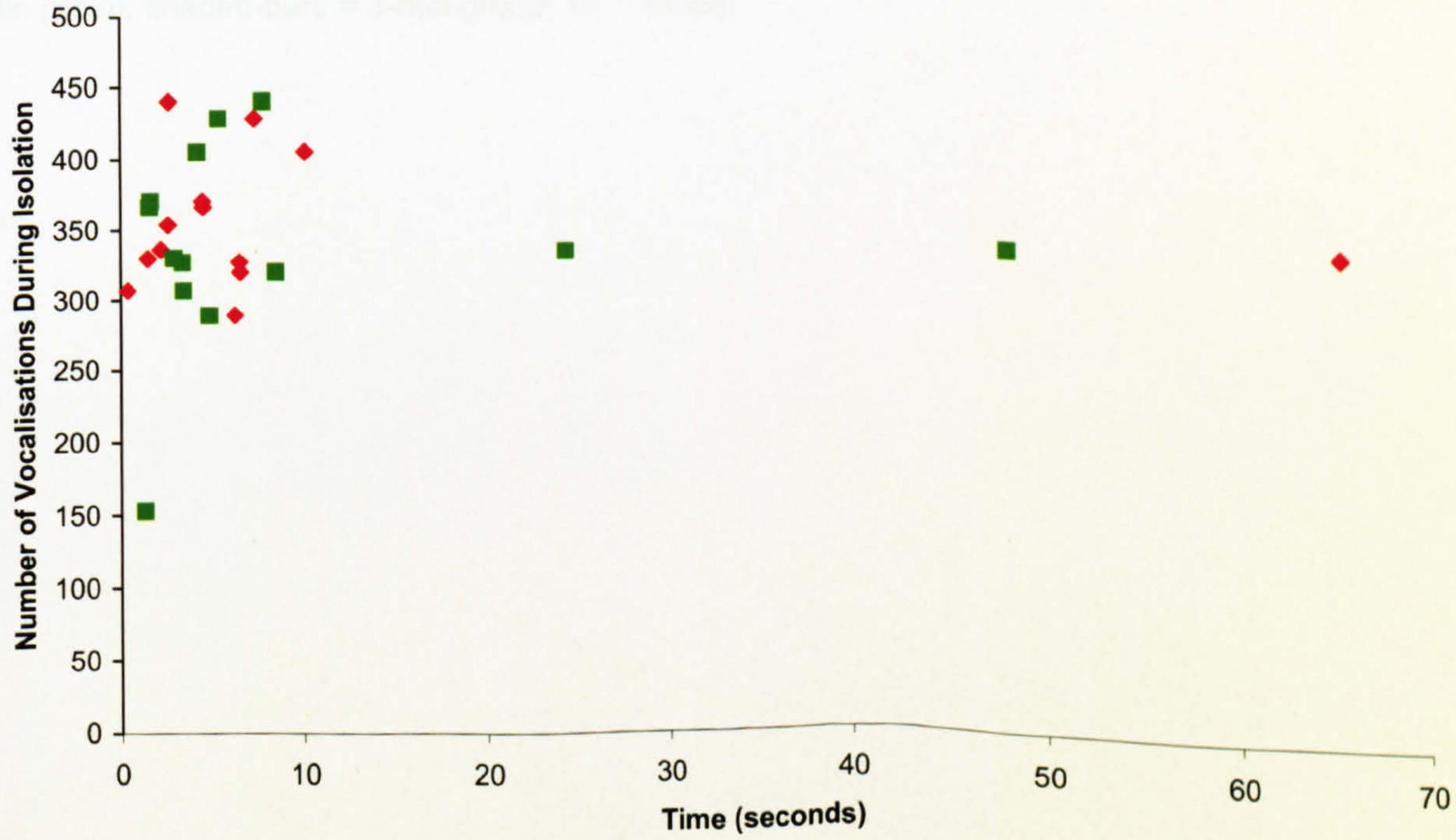

Figure 7.74 The number of vocalisations recorded from the chicks in the 3-min group during their three minute period of social isolation, against their latency to eat the first red (represented by diamonds) or green crumbs (represented by squares) in their subsequent test session.

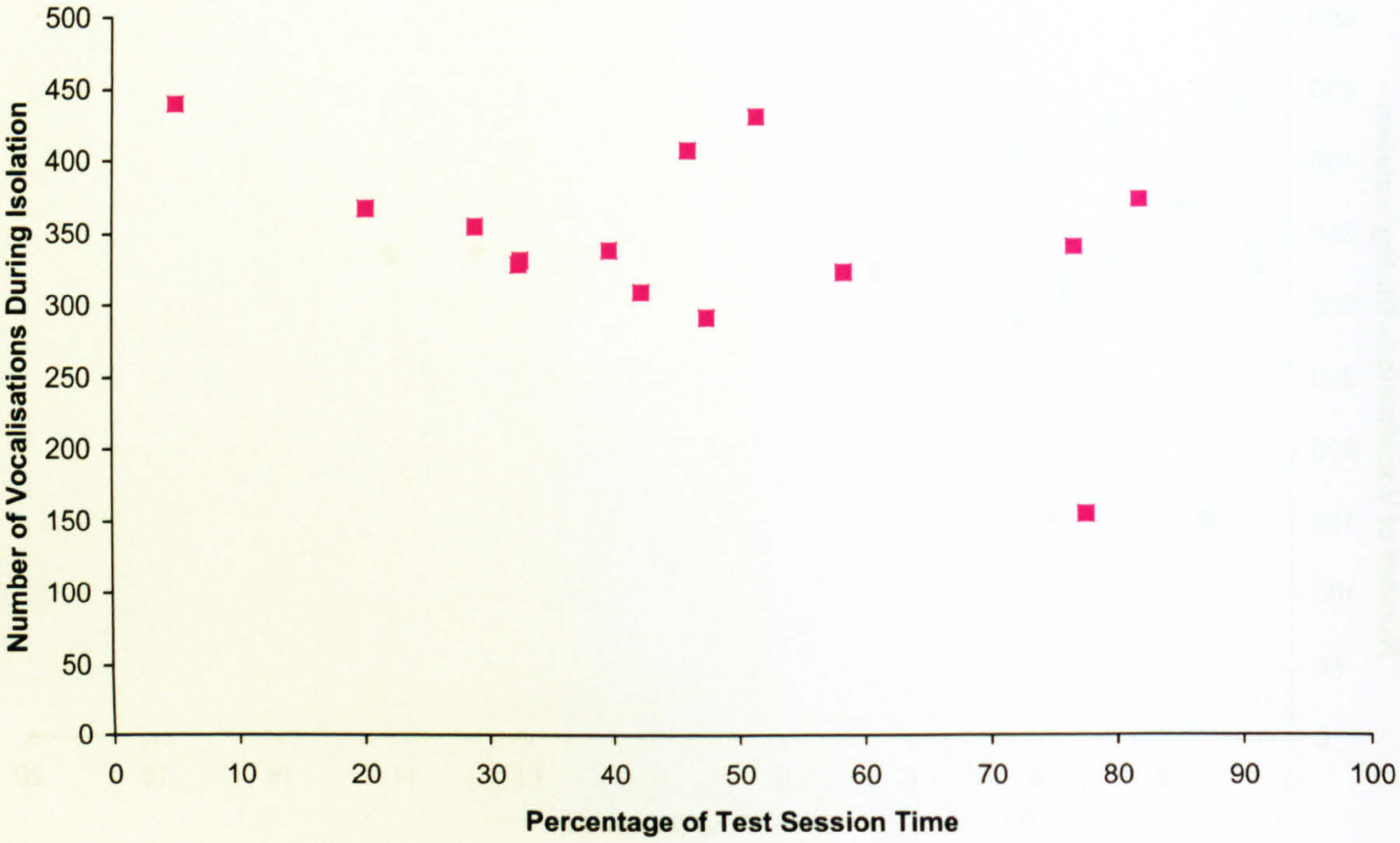

Figure 7.75 The number of vocalisations recorded from the chicks in the 3-min group during their three minute period of social isolation, against the percentage of time spent in the area of the test arena closest to the buddy chicks (Areas 1 & 2; see Figure 5.1) during their subsequent test session.

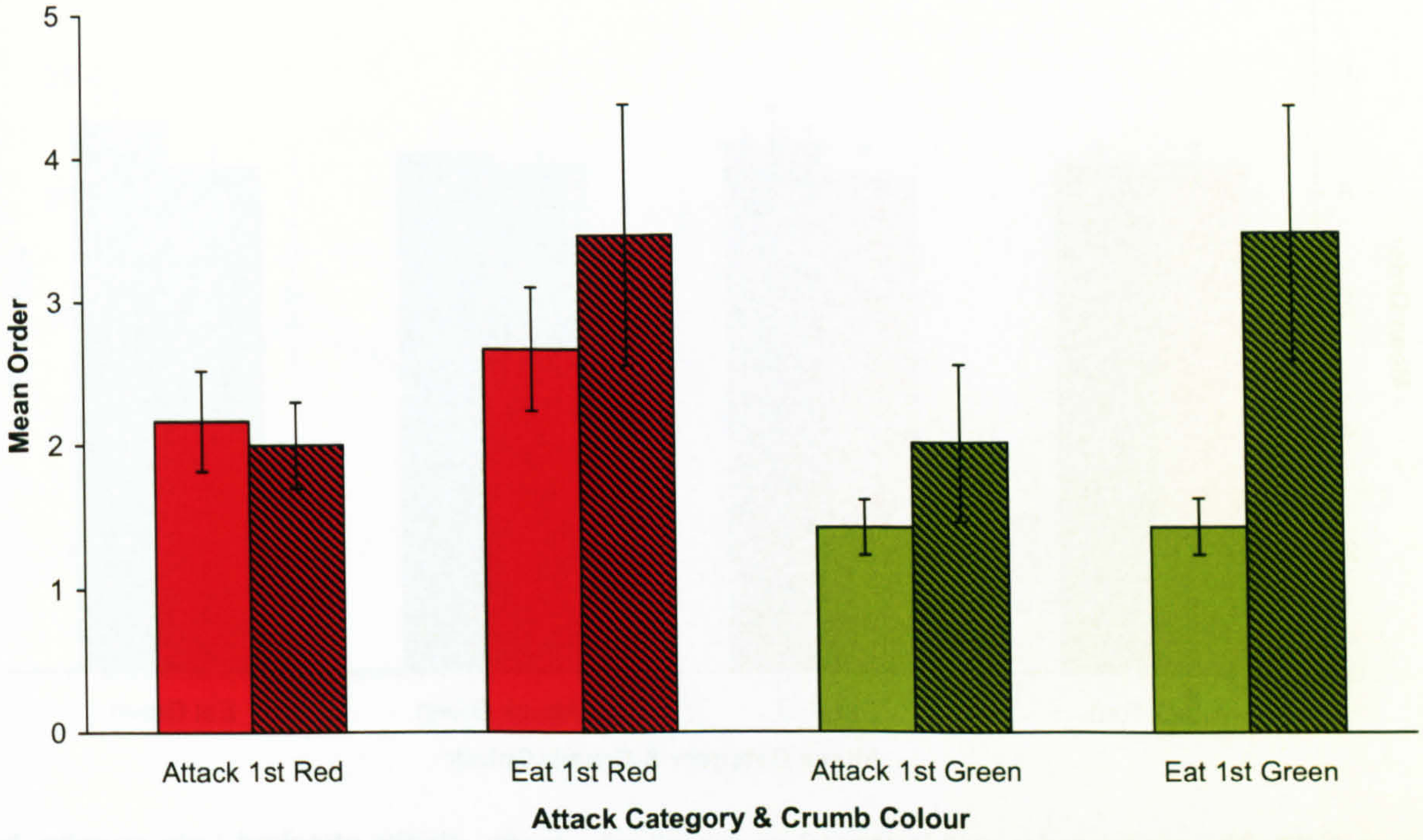

Figure 7.76 The mean order, within the 16 crumb attacks made, chicks attacked / ate the *first* crumb, by treatment group and crumb colour (NB an 'attack' is any contact between beak and crumb, regardless of whether that crumb is subsequently eaten, or pecked & rejected; for each bar, only data from chicks attacking / eating a crumb of that colour are included; non-shaded bars = 0-min group, shaded-bars = 3-min group; +/- 1 SEM).

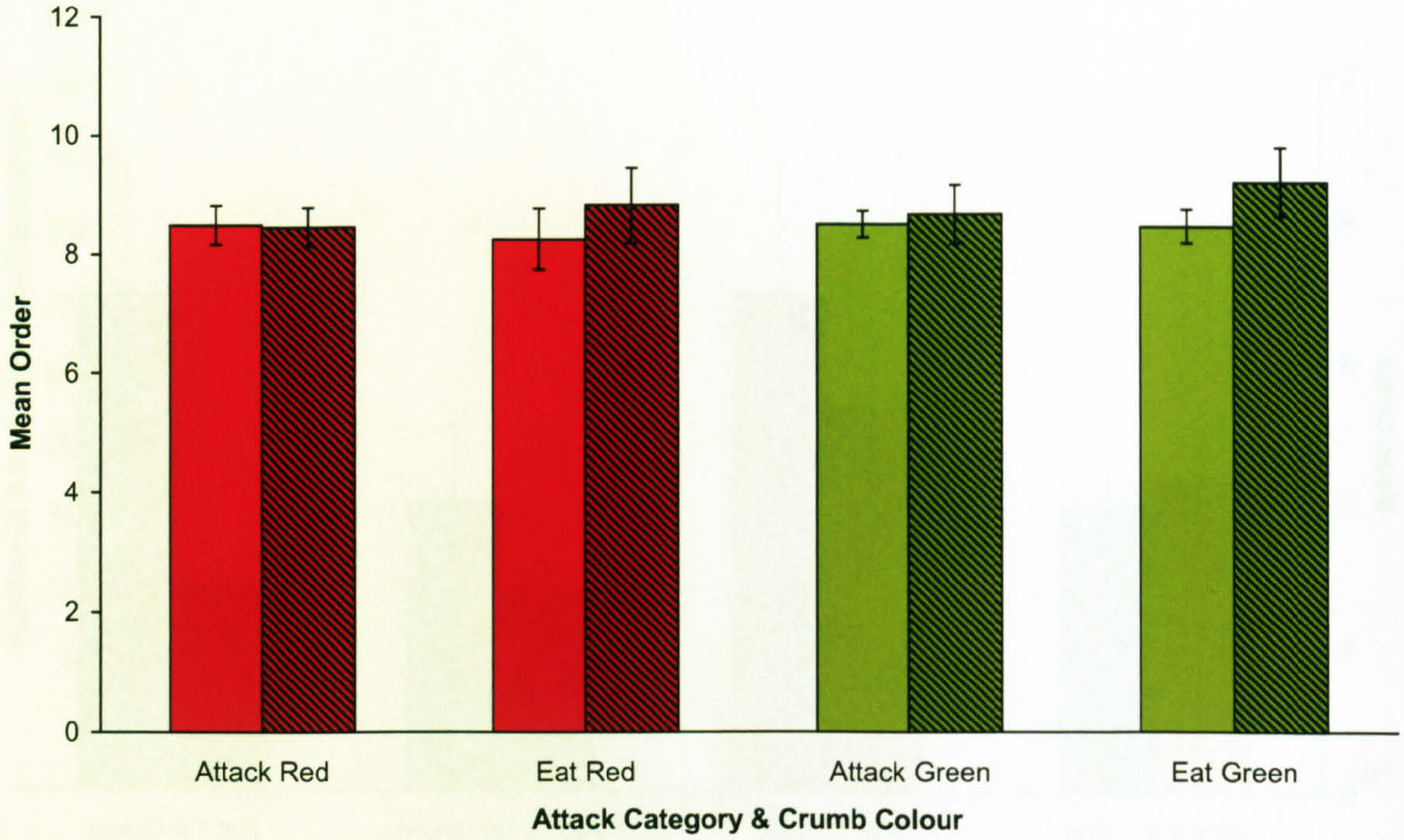

Figure 7.77 The mean order, within the 16 crumb attacks made, chicks attacked / ate crumbs, by treatment group and crumb colour (NB an 'attack' is any contact between beak and crumb, regardless of whether that crumb is subsequently eaten, or pecked & rejected; for each bar, only data from chicks attacking / eating a crumb of that colour are included; non-shaded bars = 0-min group, shaded-bars = 3-min group; +/- 1 SEM).

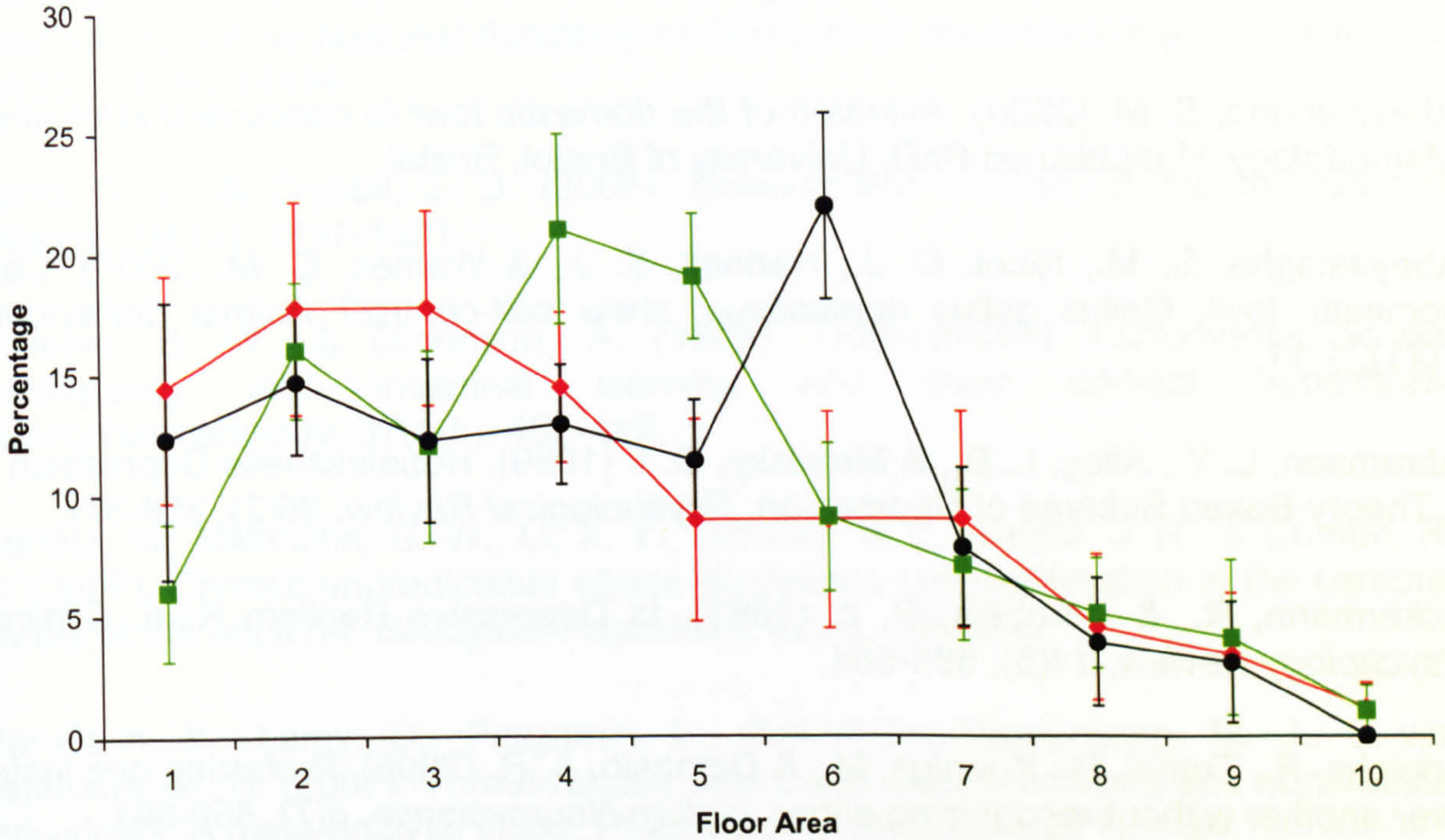

Figure 7.78 The mean percentage of red and green crumbs attacked by the chicks in the 0-min group in each of the ten Floor Areas (with Area 1 closest to the buddy chicks – see Figure 5.1), together with the mean percentage of test session time they spent in each area. (Red diamonds = red crumbs; green squares = green crumbs; black circles = time; +/- 1SEM).

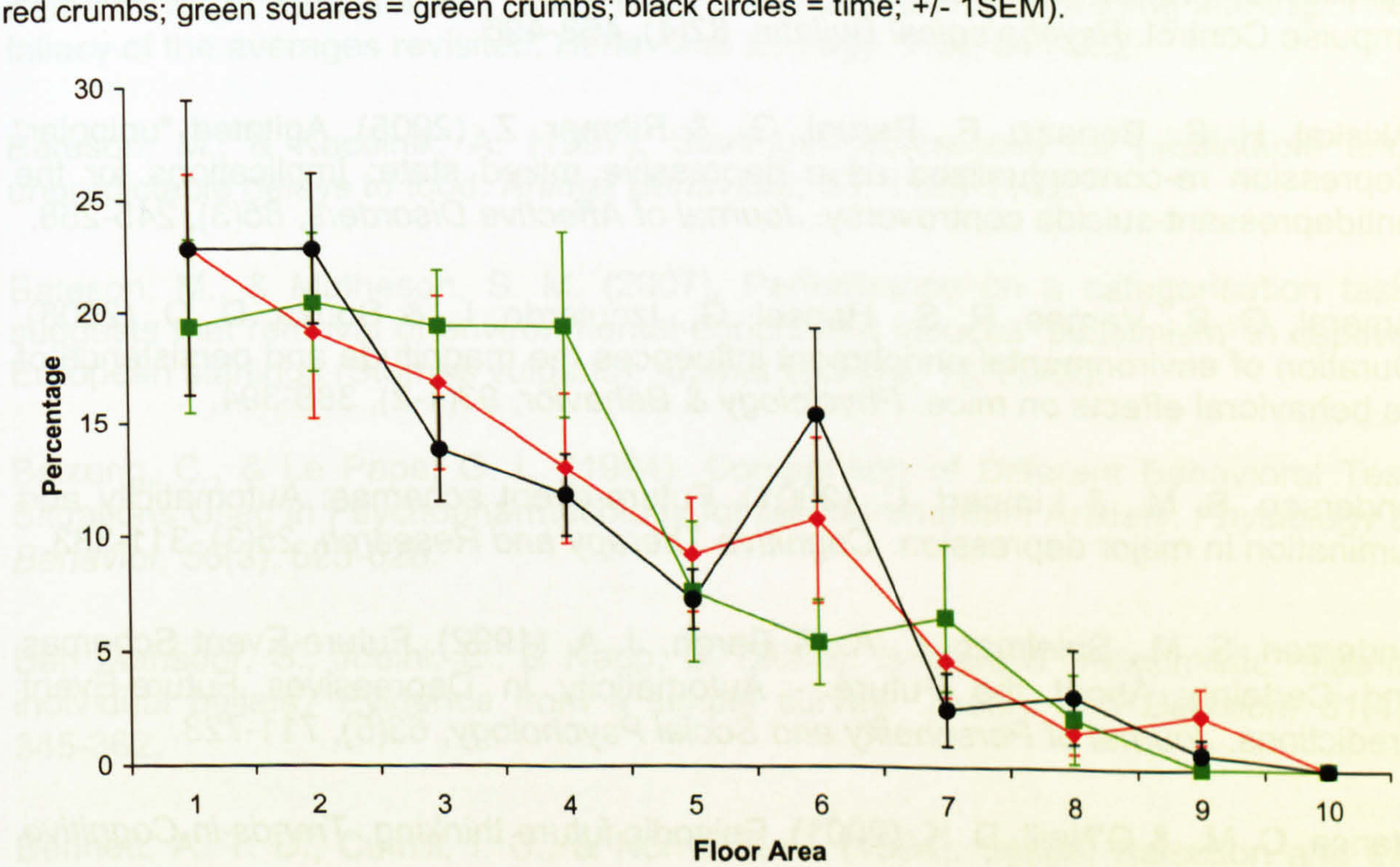

Figure 7.79 As Figure 7.78, but for the chicks in the 3-min group.

REFERENCES

Abeyesinghe, S. M. (2000). *Aversion of the domestic fowl to concurrent stressors: Methodology.* Unpublished PhD, University of Bristol, Bristol.

Abeyesinghe, S. M., Nicol, C. J., Hartnell, S. J., & Wathes, C. M. (2005). Can domestic fowl, Gallus gallus domesticus, show self-control? *Animal Behaviour, 70*(1), 1-11.

Abramson, L. Y., Alloy, L. B., & Metalsky, G. I. (1989). Hopelessness Depression - a Theory-Based Subtype of Depression. *Psychological Review, 96*(2), 358-372.

Ackermann, R., & Derubeis, R. J. (1991). Is Depressive Realism Real. *Clinical Psychology Review, 11*(5), 565-584.

Adolphs, R., Tranel, D., Koenigs, M., & Damasio, A. R. (2005). Preferring one taste over another without recognizing either. *Nature Neuroscience, 8*(7), 860-861.

Aiken, L. S., & West, S. G. (1991). *Multiple Regression: Testing and Interpreting Interactions.* London, U.K.: Sage Publications.

Ainslie, G. (1975). Specious Reward - Behavioral Theory of Impulsiveness and Impulse Control. *Psychological Bulletin, 82*(4), 463-496.

Akiskal, H. S., Benazzi, F., Perugi, G., & Rihmer, Z. (2005). Agitated "unipolar" depression re-conceptualized as a depressive mixed state: Implications for the antidepressant-suicide controversy. *Journal of Affective Disorders, 85*(3), 245-258.

Amaral, O. B., Vargas, R. S., Hansel, G., Izquierdo, I., & Souza, D. O. (2008). Duration of environmental enrichment influences the magnitude and persistence of its behavioral effects on mice. *Physiology & Behavior, 93*(1-2), 388-394.

Andersen, S. M., & Limpert, C. (2001). Future-event schemas: Automaticity and rumination in major depression. *Cognitive Therapy and Research, 25*(3), 311-333.

Andersen, S. M., Spielman, L. A., & Bargh, J. A. (1992). Future-Event Schemas and Certainty About the Future - Automaticity in Depressives Future-Event Predictions. *Journal of Personality and Social Psychology, 63*(5), 711-723.

Atance, C. M., & O'Neill, D. K. (2001). Episodic future thinking. *Trends in Cognitive Sciences, 5*(12), 533-539.

Atkinson, R. L., Atkinson, R. C., Smith, E. E., Bem, D. J., & Nolen-Hoeksema, S. (1996). *Hilgard's Introduction to Psychology.* Orlando, FL.: Harcourt Bryce.

Azzopardi, P., & Cowey, A. (1997). Is blindsight like normal, near-threshold vision? *Proceedings of the National Academy of Sciences of the United States of America, 94*(25), 14190-14194.

Babb, S. J., & Crystal, J. D. (2006). Episodic-like memory in the rat. *Current Biology, 16*(13), 1317-1321.

Balleine, B. W., & Dickinson, A. (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology, 37*(4-5), 407-419.

Banasr, M., Valentine, G. W., Li, X. Y., Gourley, S. L., Taylor, J. R., & Duman, R. S. (2007). Chronic unpredictable stress decreases cell proliferation in the cerebral cortex of the adult rat. *Biological Psychiatry, 62*(5), 496-504.

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van Ijendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. *Psychological Bulletin, 133*(1), 1-24.

Barnard, C. (2004). *Animal Behaviour: Mechanism, Development, Function and Evolution.* Harlow, Essex: Pearson Education Limited.

Bateson, M., & Kacelnik, A. (1996). Rate currencies and the foraging starling: The fallacy of the averages revisited. *Behavioral Ecology, 7*(3), 341-352.

Bateson, M., & Kacelnik, A. (1997). Starlings' preferences for predictable and unpredictable delays to food. *Animal Behaviour, 53*, 1129-1142.

Bateson, M., & Matheson, S. M. (2007). Performance on a categorisation task suggests that removal of environmental enrichment induces 'pessimism' in captive European starlings (Sturnus vulgaris). *Animal Welfare, 16*, 33-36.

Belzung, C., & Le Pape, G. L. (1994). Comparison of Different Behavioral Test Situations Used in Psychopharmacology for Measurement of Anxiety. *Physiology & Behavior, 56*(3), 623-628.

Ben Mansour, S., Jouini, E., & Napp, C. (2006). Is there a "Pessimistic" bias in individual beliefs? Evidence from a simple survey. *Theory and Decision, 61*(4), 345-362.

Bennett, A. T. D., Cuthill, I. C., & Norris, K. J. (1994). Sexual Selection and the Mismeasure of Color. *American Naturalist, 144*(5), 848-860.

Berlin, H. A., Rolls, E. T., & Kischka, U. (2004). Impulsivity, time perception, emotion and reinforcement sensitivity in patients with orbitofrontal cortex lesions. *Brain, 127*, 1108-1126.

Berridge, K. C. (2003). Comparing the Emotional Brains of Humans and Other Animals. In R. J. Davidson, K. R. Scherer & H. H. Goldsmith (Eds.), *Handbook of Affective Sciences*. Oxford: Oxford University Press.

Biringer, E., Lundervold, A., Stordal, K., Mykletun, A., Egeland, J., Bottlender, R., et al. (2005). Executive function improvement upon remission of recurrent unipolar depression. *European Archives of Psychiatry and Clinical Neuroscience, 255*(6), 373-380.

Bishop, S. J. (2007). Neurocognitive mechanisms of anxiety: an integrative account. *Trends in Cognitive Sciences, 11*(7), 307-316.

Bisson, M. A. S., & Sears, C. R. (2007). The effect of depressed mood on the interpretation of ambiguity, with and without negative mood induction. *Cognition & Emotion, 21*(3), 614-645.

Bizarro, L., & Stolerman, I. P. (2003). Attentional effects of nicotine and amphetamine in rats at different levels of motivation. *Psychopharmacology, 170*(3), 271-277.

Blackmore, S. (2003). *Consciousness: An Introduction*. Oxford: Oxford University Press.

Blanchard, D. C., Hynd, A. L., Minke, K. A., Minemoto, T., & Blanchard, R. J. (2001). Human defensive behaviors to threat scenarios show parallels to fear- and anxiety-related defense patterns of non-human mammals. *Neuroscience and Biobehavioral Reviews, 25*(7-8), 761-770.

Blaney, P. H. (1986). Affect and Memory - a Review. *Psychological Bulletin, 99*(2), 229-246.

Blewett, A. E. (1992). Abnormal Subjective Time Experience in Depression. *British Journal of Psychiatry, 161*, 195-200.

Bliss, C. I. (1934). The method of probits. *Science, 79*, 38-39.

Block, N. (1995). On A Confusion About A Function Of Consciousness. *Behavioral And Brain Sciences, 18*(2), 227-247.

Bokkers, E. A. M., Koene, P., Rodenburg, T. B., Zimmerman, P. H., & Spruijt, B. M. (2004). Working for food under conditions of varying motivation in broilers. *Animal Behaviour, 68*, 105-113.

Bradley, B. P., Hogg, K., White, J., Groom, C., & de Bono, J. (1999). Attentional bias for emotional faces in generalized anxiety disorder. *British Journal of Clinical Psychology, 38*, 267-278.

Bradley, B. P., Mogg, K., & Williams, R. (1995). Implicit and Explicit Memory for Emotion-Congruent Information in Clinical Depression and Anxiety. *Behaviour Research and Therapy, 33*(7), 755-770.

Brake, J., Keeley, T. P., & Jones, R. B. (1994). Effect of Age and Presence of Perches During Rearing on Tonic Immobility Fear Reactions of Broiler Breeder Pullets. *Poultry Science, 73*(9), 1470-1474.

Bronson, G. W. (1968). Fear of Novelty. *Psychological Bulletin, 69*(5), 350-358.

Broom, D. M., & Johnson, K. G. (1993). *Stress and Animal Welfare*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Bschor, T., Ising, M., Bauer, M., Lewitzka, U., Skerstupeit, M., Muller-Oerlinghausen, B., et al. (2004). Time experience and time judgment in major depression, mania and healthy subjects. A controlled study of 93 subjects. *Acta Psychiatrica Scandinavica, 109*(3), 222-229.

Burman, O. H. P., Abou Ismail, U., Nicol, C. J., Day, M., Owen, D., Bailey, M., et al. (2006). *A Multidisciplinary Study of the Long-Term Effects of Environmental Enrichment on Laboratory Rat Welfare.* Paper presented at the International Congress of the ISAE, Bristol, UK.

Burman, O. H. P., Parker, R. M. A., Paul, E. S., & Mendl, M. (2008a). Sensitivity to reward loss as an indicator of animal emotion and welfare. *Biology Letters, 4*(4), 330-333.

Burman, O. H. P., Parker, R. M. A., Paul, E. S., & Mendl, M. (2008b). A spatial judgement task to detect background emotional state in laboratory rats (Rattus norvegicus). . *Animal Behaviour, 76*(3), 801-809.

Burt, D. B., Zembar, M. J., & Niederehe, G. (1995). Depression and Memory Impairment - a Metaanalysis of the Association, Its Pattern, and Specificity. *Psychological Bulletin, 117*(2), 285-305.

Butler, A. B., & Cotterill, R. M. J. (2006). Mammalian and avian neuroanatomy and the question of consciousness in birds. *Biological Bulletin, 211*(2), 106-127.

Butler, G., & Mathews, A. (1983). Cognitive-Processes in Anxiety. *Advances in Behaviour Research and Therapy, 5*(1), 51-62.

Byrne, A., & MacLeod, A. K. (1997). Attributions and accessibility of explanations for future events in anxiety and depression. *British Journal of Clinical Psychology, 36*, 505-520.

Byrne, R. W., & Bates, L. A. (2006). Why are animals cognitive? *Current Biology, 16*(12), R445-R448.

Cabeza, R., & Kingstone, A. (Eds.). (2006). *Handbook of functional neuroimaging of cognition.* Cambridge, Mass.: MIT Press.

Cabib, S. (1997). What is mild in mild stress? *Psychopharmacology, 134*(4), 344-346.

Cabib, S., & PuglisiAllegra, S. (1996). Stress, depression and the mesolimbic dopamine system. *Psychopharmacology, 128*(4), 331-342.

Calvo, M. G., & Castillo, M. D. (1998). Predictive inferences take time to develop. *Psychological Research-Psychologische Forschung, 61*(4), 249-260.

Calvo, M. G., Eysenck, M. W., & Estevez, A. (1994). Ego-Threat Interpretive Bias in Test Anxiety - Online Inferences. *Cognition & Emotion, 8*(2), 127-146.

Carobrez, A. P., & Bertoglio, L. J. (2005). Ethological and temporal analyses of anxiety-like behavior: The elevated plus-maze model 20 years on. *Neuroscience and Biobehavioral Reviews, 29*(8), 1193-1205.

Carver, C. S. (2001). Affect and the functional bases of behavior: On the dimensional structure of affective experience. *Personality and Social Psychology Review, 5*(4), 345-356.

Castiello, U., Paulignan, Y., & Jeannerod, M. (1991). Temporal Dissociation of Motor-Responses and Subjective Awareness - a Study in Normal Subjects. *Brain, 114*, 2639-2655.

Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies, 2*(3), 200-219.

Chalmers, D. J. (1996). *The Conscious Mind.* Oxford: Oxford University Press.

Changizi, M. A., & Hall, W. G. (2001). Thirst modulates a perception. *Perception, 30*(12), 1489-1497.

Church, R. M., & Deluty, M. Z. (1977). Bisection of temporal intervals. *Journal Of Experimental Psychology: Animal Behavior Processes, 3*, 216-228.

Church, R. M., Miller, K. D., Meck, W. H., & Gibbon, J. (1991). Symmetrical And Asymmetrical Sources Of Variance In Temporal Generalization. *Animal Learning & Behavior, 19*(3), 207-214.

Clark, D. M., & Teasdale, J. D. (1982). Diurnal-Variation in Clinical Depression and Accessibility of Memories of Positive and Negative Experiences. *Journal of Abnormal Psychology, 91*(2), 87-95.

Clark, L. A., & Watson, D. (1991). Tripartite Model of Anxiety and Depression - Psychometric Evidence and Taxonomic Implications. *Journal of Abnormal Psychology, 100*(3), 316-336.

Clayton, N. S., Salwiczek, L. H., & Dickinson, A. (2007). Episodic memory. *Current Biology, 17*(6), R189-R191.

Clore, G. L., & Huntsinger, J. R. (2007). How emotions inform judgment and regulate thought. *Trends in Cognitive Sciences, 11*(9), 393-399.

Coles, M. E., & Heimberg, R. G. (2002). Memory biases in the anxiety disorders: Current status. *Clinical Psychology Review, 22*(4), 587-627.

Connor, C. E., Egeth, H. E., & Yantis, S. (2004). Visual attention: Bottom-up versus top-down. *Current Biology, 14*(19), R850-R852.

Coricelli, G., Dolan, R. J., & Sirigu, A. (2007). Brain, emotion and decision making: the paradigmatic example of regret. *Trends in Cognitive Sciences, 11*(6), 258-265.

Cott, H. B. (1940). *Adaptive Coloration in Animals*. London: Methuen & Co. Ltd.

Cowey, A., & Stoerig, P. (1995). Blindsight in Monkeys. *Nature, 373*(6511), 247-249.

Cropley, M., MacLeod, A. K., & Tata, P. (2000). Memory retrieval and subjective probability judgements in control and depressed participants. *Clinical Psychology & Psychotherapy, 7*(5), 367-378.

Cruz, A. P. M., Pinheiro, G., Alves, S. H., Ferreira, G., Mendes, M., Faria, L., et al. (2005). Behavioral effects of systemically administered MK-212 are prevented by ritanserin microinfusion into the basolateral amygdala of rats exposed to the elevated plus-maze. *Psychopharmacology, 182*(3), 345-354.

Crystal, J. D. (1999). Systematic nonlinearities in the perception of temporal intervals. *Journal of Experimental Psychology-Animal Behavior Processes, 25*(1), 3-17.

Crystal, J. D. (2001). Nonlinear time perception. *Behavioural Processes, 55*(1), 35-49.

Cuthill, I. C., Bennett, A. T. D., Partridge, J. C., & Maier, E. J. (1999). Plumage Reflectance and the Objective Assessment of Avian Sexual Dichromatism. *American Naturalist, 153*(2), 183-200.

D'Aquila, P. S., Brain, P., & Willner, P. (1994). Effects of Chronic Mild Stress on Performance in Behavioral-Tests Relevant to Anxiety and Depression. *Physiology & Behavior, 56*(5), 861-867.

D'Esposito, M., & Postle, B. R. (2000). Neural Correlates of Processes Contributing to Working-Memory Function: Evidence from Neuropsychological and Pharmacological Studies. In S. Monsell & J. Driver (Eds.), *Control of Cognitive Processes: Attention and Performance XVIII*. Cambridge, Massachusetts: The MIT Press.

D'Haenen, H., & Andrews, J. S. (2000). Animal models of affective disorders. *Neuroscience Research Communications, 26*(3), 289-300.

Dalgleish, T. (2003). Information Processing Approaches to Emotion. In R. J. Davidson, K. R. Scherer & H. H. Goldsmith (Eds.), *Handbook of Affective Sciences*. Oxford: Oxford University Press.

Dalgleish, T., & Watts, F. N. (1990). Biases of Attention and Memory in Disorders of Anxiety and Depression. *Clinical Psychology Review, 10*(5), 589-604.

Damasio, A. R. (1994). *Descartes' Error: Emotion, Reason and the Human Brain*. New York: Avon.

Dawkins, M. S. (1990). From An Animals Point Of View - Motivation, Fitness, And Animal-Welfare. *Behavioral And Brain Sciences, 13*(1), 1-&.

Dawkins, M. S. (2001). Who needs consciousness? *Animal Welfare, 10*, S19-S29.

Den Hartog, H. M., Derix, M. M. A., van Bemmel, A. L., Kremer, B., & Jolles, J. (2003). Cognitive functioning in young and middle-aged unmedicated out-patients with major depression: testing the effort and cognitive speed hypotheses. *Psychological Medicine, 33*(8), 1443-1451.

Dent, J., & Teasdale, J. D. (1988). Negative Cognition and the Persistence of Depression. *Journal of Abnormal Psychology, 97*(1), 29-34.

Descartes, R. (1968). *Discourse on Method and the Meditations*. Harmondsworth, Middlesex: Penguin.

Diagnostic and Statistical Manual of Mental Disorders IV. (1994).). Washington DC: American Psychiatric Association.

Dickinson, A., & Balleine, B. (1995). Motivational Control of Instrumental Action. *Current Directions in Psychological Science, 4*(5), 162-167.

Dickinson, A., Balleine, B., Watt, A., Gonzalez, F., & Boakes, R. A. (1995). Motivational Control after Extended Instrumental Training. *Animal Learning & Behavior, 23*(2), 197-206.

Dickinson, A., Campos, J., Varga, Z. I., & Balleine, B. (1996). Bidirectional instrumental conditioning. *Quarterly Journal of Experimental Psychology Section B-Comparative and Physiological Psychology, 49*(4), 289-306.

Droit-Volet, S., & Meck, W. H. (2007). How emotions colour our perception of time. *Trends in Cognitive Sciences, 11*(12), 504-513.

Duncan, S., & Barrett, L. F. (2007). Affect is a form of cognition: A neurobiological analysis. *Cognition & Emotion, 21*(6), 1184-1211.

Ebmeier, K. P., Donaghey, C., & Steele, J. D. (2006). Recent developments and current controversies in depression. *Lancet, 367*(9505), 153-167.

Echevarria, D. J., Brewer, A., Burk, J. A., Brown, S. N., Manuzon, H., & Robinson, J. K. (2005). Construct validity of an operant signal detection task for rats. *Behavioural Brain Research, 157*(2), 283-290.

Edelman, D. B., Baars, B. J., & Seth, A. K. (2005). Identifying hallmarks of consciousness in non-mammalian species. *Consciousness and Cognition, 14*(1), 169-187.

Eder, A. B., Hommel, B., & De Houwer, J. (2007). How distinctive is affective processing? On the implications of using cognitive paradigms to study affect and emotion. *Cognition & Emotion, 21*(6), 1137-1154.

Edmunds, M. (1974). *Defence in Animals*. Harlow, Essex: Longman Group Ltd.

Ekman, P. (1999). Basic Emotions. In T. Dalgleish & M. Power (Eds.), *Handbook of Cognition and Emotion*. Chichester: John Wiley & Sons Ltd.

Emerson, J. D. (1991). Introduction to Transformation. In D. C. Hoaglin, F. Mosteller & J. W. Tukey (Eds.), *Fundamentals of Exploratory Analysis of Variance*. New York: John Wiley & Sons, Inc.

Evenden, J. L. (1999). Varieties of impulsivity. *Psychopharmacology, 146*(4), 348-361.

Eysenck, M., & Keane, M. (2005). *Cognitive Psychology*. Hove, East Sussex: Psychology Press Ltd.

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and Cognitive Performance: Attentional Control Theory. *Emotion, 7*(2), 336-353.

Eysenck, M. W., Mogg, K., May, J., Richards, A., & Mathews, A. (1991). Bias in Interpretation of Ambiguous Sentences Related to Threat in Anxiety. *Journal of Abnormal Psychology, 100*(2), 144-150.

Eysenck, M. W., Payne, S., & Santos, R. (2006). Anxiety and depression: Past, present, and future events. *Cognition & Emotion, 20*(2), 274-294.

Fawcett, J., Clark, D. C., Scheftner, W. A., & Gibbons, R. D. (1983). Assessing Anhedonia In Psychiatric-Patients - The Pleasure Scale. *Archives Of General Psychiatry, 40*(1), 79-84.

Feltenstein, M. W., Ford, N. G., Freeman, K. B., & Sufka, K. J. (2002). Dissociation of stress behaviors in the chick social-separation-stress procedure. *Physiology & Behavior, 75*(5), 675-679.

Feltenstein, M. W., Warnick, J. E., Guth, A. N., & Sufka, K. J. (2004). The chick separation stress paradigm: a validation study. *Pharmacology Biochemistry and Behavior, 77*(2), 221-226.

Ferguson, S. A., & Paule, M. G. (1997). Progressive ratio performance varies with body weight in rats. *Behavioural Processes, 40*(2), 177-182.

Field, A. (2000). *Discovering Statistics Using SPSS For Windows*. London, U.K.: SAGE Publications Ltd.

File, S. E. (2001). Factors controlling measures of anxiety and responses to novelty in the mouse. *Behavioural Brain Research, 125*(1-2), 151-157.

Finney, D. J. (1971). *Probit Analysis* (3rd. ed.). Cambridge: Cambridge University Press.

Finucane, M. L., Alhakami, A., Slovic, P., & Johnson, S. M. (2000). The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making, 13*(1), 1-17.

Flaherty, C. F. (1996). *Incentive Relativity*. Cambridge: Cambridge University Press.

Forbes, N. F., Stewart, C. A., Matthews, K., & Reid, I. C. (1996). Chronic mild stress and sucrose consumption: Validity as a model of depression. *Physiology & Behavior, 60*(6), 1481-1484.

Forgas, J. P. (1995). Mood and Judgment - the Affect Infusion Model (Aim). *Psychological Bulletin, 117*(1), 39-66.

Forgas, J. P. (2003). Affective Influences on Attitudes and Judgments. In R. J. Davidson, K. R. Scherer & H. H. Goldsmith (Eds.), *Handbook of Affective Sciences*. Oxford: Oxford University Press.

Fox, C., Merali, Z., & Harrison, C. (2006). Therapeutic and protective effect of environmental enrichment against psychogenic and neurogenic stress. *Behavioural Brain Research, 175*(1), 1-8.

Francia, N., Cirulli, F., Chiarotti, F., Antonelli, A., Aloe, L., & Alleva, E. (2006). Spatial memory deficits in middle-aged mice correlate with lower exploratory activity and a subordinate status: role of hippocampal neurotrophins. *European Journal of Neuroscience, 23*(3), 711-728.

Frijda, N. H. (1994). Varieties of Affect: Emotions and Episodes, Moods, and Sentiments. In P. Ekman & R. J. Davidson (Eds.), *The Nature of Emotion: Fundamental Questions*. Oxford: Oxford University Press.

Gamberale-Stille, G., Hall, K. S. S., & Tullberg, B. S. (2007). Signals of profitability? Food colour preferences in migrating juvenile blackcaps differ for fruits and insects. *Evolutionary Ecology, 21*(1), 99-108.

Gamberale-Stille, G., & Tullberg, B. S. (2001). Fruit or aposematic insect? Context-dependent colour preferences in domestic chicks. *Proceedings of the Royal Society of London Series B-Biological Sciences, 268*(1485), 2525-2529.

Garcia, A. M. B., Martinez, R., Brandao, M. L., & Morato, S. (2005). Effects of apomorphine on rat behavior in the elevated plus-maze. *Physiology & Behavior, 85*(4), 440-447.

Ghirlanda, S., & Enquist, M. (2003). A century of generalization. *Animal Behaviour, 66*, 15-36.

Goldstein, H. (1986). Multilevel mixed linear model analysis using iterative generalized least squares. *Biometrika, 73*(1), 43-56.

Goldstein, H. (2003). *Multilevel Statistical Models* (3rd ed.). London: Edward Arnold.

Goldstein, H., Healy, M. J. R., & Rasbash, J. (1994). Multilevel Time-Series Models with Applications to Repeated-Measures Data. *Statistics in Medicine, 13*(16), 1643-1655.

Gotlib, I. H., & Krasnoperova, E. (1998). Biased information processing as a vulnerability factor for depression. *Behavior Therapy, 29*(4), 603-617.

Gouirand, A. M., & Matuszewich, L. (2005). The effects of chronic unpredictable stress on male rats in the water maze. *Physiology & Behavior, 86*(1-2), 21-31.

Grafen, A., & Hails, R. (2002). *Modern Statistics for the Life Sciences*. Oxford: Oxford University Press.

Gray, J. A. (1982). *The Neuropsychology of Anxiety: An Enquiry into the Functions of the Septo-hippocampal System*. Oxford: Oxford University Press.

Gray, J. A. (1999). Cognition, Emotion, Conscious Experience and the Brain. In T. Dagleish & M. Power (Eds.), *Handbook of Cognition and Emotion*. Chichester: John Wiley & Sons Ltd.

Gregory, R. L. (1998). *Eye and Brain: The Psychology of Seeing* (Fifth Edition ed.). Oxford: Oxford University Press.

Gronli, J., Murison, R., Fiske, E., Bjorvatn, B., Sorensen, E., Portas, C. M., et al. (2005). Effects of chronic mild stress on sexual behavior, locomotor activity and consumption of sucrose and saccharine solutions. *Physiology & Behavior, 84*(4), 571-577.

Gualtieri, C. T., Johnson, L. G., & Benedict, K. B. (2006). Neurocognition in depression: Patients on and off medication versus healthy comparison subjects. *Journal of Neuropsychiatry and Clinical Neurosciences, 18*(2), 217-225.

Guilford, T., & Dawkins, M. S. (1991). Receiver psychology and the evolution of animal signals. *Animal Behaviour, 42*(1), 1-14.

Haller, J., & Halasz, J. (1999). Mild social stress abolishes the effects of isolation on anxiety and chlordiazepoxide reactivity. *Psychopharmacology, 144*(4), 311-315.

Haller, J., & Halasz, J. (2000). Anxiolytic effects of repeated victory in male Wistar rats. *Aggressive Behavior, 26*(3), 257-261.

Ham, A. D., & Osorio, D. (2007). Colour Preferences and Colour Vision in Poultry Chicks. *Proceedings of the Royal Society B-Biological Sciences, 274*, 1941-1948.

Harding, E. J. (2002). *Novel methods for assessing mental states and animal welfare*. University of Bristol, Bristol.

Harding, E. J., Paul, E. S., & Mendl, M. (2004). Animal behavior - Cognitive bias and affective state. *Nature, 427*(6972), 312-312.

Harkin, A., Houlihan, D. D., & Kelly, J. P. (2002). Reduction in preference for saccharin by repeated unpredictable stress in mice and its prevention by imipramine. *Journal of Psychopharmacology, 16*(2), 115-123.

Harris, R. B. S., Zhou, J., Youngblood, B. D., Rybkin, II, Smagin, G. N., & Ryan, D. H. (1998). Effect of repeated stress on body weight and body composition of rats fed low- and high-fat diets. *American Journal of Physiology-Regulatory Integrative and Comparative Physiology, 44*(6), R1928-R1938.

Harris, R. B. S., Zhou, J., Youngblood, B. D., Smagin, G. N., & Ryan, D. H. (1997). Failure to change exploration or saccharin preference in rats exposed to chronic mild stress. *Physiology & Behavior, 63*(1), 91-100.

Hastie, R. (2001). Problems for judgment and decision making. *Annual Review of Psychology, 52*, 653-683.

Haug, M., & Whalen, R. E. (Eds.). (1999). *Animal Models of Human Emotion and Cognition*. Washington DC.: American Psychological Association.

Hebets, E. A., & Papaj, D. R. (2005). Complex signal function: developing a framework of testable hypotheses. *Behavioral Ecology and Sociobiology, 57*(3), 197-214.

Heisler, L. K., Chu, H. M., Brennan, T. J., Danao, J. A., Bajwa, P., Parsons, L. H., et al. (1998). Elevated anxiety and antidepressant-like responses in serotonin 5-HT1A receptor mutant mice. *Proceedings of the National Academy of Sciences of the United States of America, 95*(25), 15049-15054.

Hoch, S. J., & Loewenstein, G. F. (1991). Time-Inconsistent Preferences and Consumer Self-Control. *Journal of Consumer Research, 17*(4), 492-507.

Hockey, G. R. J., Maule, A. J., Clough, P. J., & Bdzola, L. (2000). Effects of negative mood states on risk in everyday decision making. *Cognition & Emotion, 14*(6), 823-855.

Hogg, S. (1996). A review of the validity and variability of the elevated plus-maze as an animal model of anxiety. *Pharmacology Biochemistry and Behavior, 54*(1), 21-30.

Holmes, A., Parmigiani, S., Ferrari, P. F., Palanza, P., & Rodgers, R. J. (2000). Behavioral profile of wild mice in the elevated plus-maze test for anxiety. *Physiology & Behavior, 71*(5), 509-516.

Isen, A. M. (1999). Positive Affect. In T. Dagleish & M. Power (Eds.), *Handbook of Cognition and Emotion* (pp. 521-539). Chichester, England: John Wiley & Sons Ltd.

Isen, A. M., Nygren, T. E., & Ashby, F. G. (1988). Influence of Positive Affect on the Subjective Utility of Gains and Losses - It Is Just Not Worth the Risk. *Journal of Personality and Social Psychology, 55*(5), 710-717.

Ito, M. (1985). Choice and Amount of Reinforcement in Rats. *Learning and Motivation, 16*(1), 95-108.

Jetz, W., Rowe, C., & Guilford, T. (2001). Non-warning odors trigger innate color aversions - as long as they are novel. *Behavioral Ecology, 12*(2), 134-139.

Johnson, E. J., & Tversky, A. (1983). Affect, Generalization, and the Perception of Risk. *Journal of Personality and Social Psychology, 45*(1), 20-31.

Jones, G. H., Marsden, C. A., & Robbins, T. W. (1990). Increased Sensitivity to Amphetamine and Reward-Related Stimuli Following Social-Isolation in Rats - Possible Disruption of Dopamine-Dependent Mechanisms of the Nucleus-Accumbens. *Psychopharmacology, 102*(3), 364-372.

Jones, K. A., Krebs, J. R., & Whittingham, M. J. (2007). Vigilance in the third dimension: head movement not scan duration varies in response to different predator models. *Animal Behaviour, 74*, 1181-1187.

Jones, R. B. (1982). Effects of Early Environmental Enrichment Upon Open-Field Behavior and Timidity in the Domestic Chick. *Developmental Psychobiology, 15*(2), 105-111.

Jones, R. B. (1996). Fear and adaptability in poultry: Insights, implications and imperatives. *Worlds Poultry Science Journal, 52*(2), 131-174.

Jones, R. B., Blokhuis, H. J., de Jong, I. C., Keeling, L. J., McAdie, T. M., & Preisinger, R. (2004). Feather pecking in poultry: the application of science in a search for practical solutions. *Animal Welfare, 13*, S215-S219.

Jones, R. B., & Williams, J. B. (1992). Responses of Pair-Housed Male and Female Domestic Chicks to the Removal of a Companion. *Applied Animal Behaviour Science, 32*(4), 375-380.

Jones, T. M., & Elgar, M. A. (2004). The role of male age, sperm age and mating history on fecundity and fertilization success in the hide beetle. *Proceedings of the Royal Society of London Series B-Biological Sciences, 271*(1545), 1311-1318.

Joron, M. (2003). Aposematic Coloration. In V. H. Resh & R. T. Cardé (Eds.), *Encyclopedia of Insects*. San Diego, California: Academic Press.

Kacelnik, A., & Bateson, M. (1996). Risky theories - The effects of variance on foraging decisions. *American Zoologist, 36*(4), 402-434.

Kacelnik, A., & Bateson, M. (1997). Risk-sensitivity: crossroads for theories of decision-making. *Trends in Cognitive Sciences, 1*(8), 304-309.

Kacelnik, A., & Marsh, B. (2002). Cost can increase preference in starlings. *Animal Behaviour, 63*, 245-250.

Kagan, L. J., MacLeod, A. K., & Pote, H. L. (2004). Accessibility of causal explanations for future positive and negative events in adolescents with anxiety and depression. *Clinical Psychology & Psychotherapy, 11*(3), 177-186.

Kahneman, D., & Tversky, A. (1982). The simulation heuristic. In D. Kahneman, P. Slovic & A. Tversky (Eds.), *Judgement under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.

Kaiser, S., Unger, J., Kiefer, M., Markela, J., Mundt, C., & Weisbrod, M. (2003). Executive control deficit in depression: event-related potentials in a Go/Nogo task. *Psychiatry Research-Neuroimaging, 122*(3), 169-184.

Kalb, R., Dorner, M., & Kalb, S. (2006). Opposite effects of depression and antidepressants on processing speed and error rate. *Progress in Neuro-Psychopharmacology & Biological Psychiatry, 30*(2), 244-250.

Kalueff, A. V., & Tuohimaa, P. (2004). Experimental modeling of anxiety and depression. *Acta Neurobiologiae Experimentalis, 64*(4), 439-448.

Keilp, J. G., Gorlyn, M., Oquendo, M. A., Burke, A. K., & Mann, J. J. (2008). Attention deficit in depressed suicide attempters. *Psychiatry Research, 159*(1-2), 7-17.

Kelley, A. E. (1993). Locomotor Activity and Exploration. In A. Sahgal (Ed.), *Behavioural Neuroscience: A Practical Approach*. Oxford: Oxford University Press.

Klasing, K. C. (2005). Poultry nutrition: A comparative approach. *Journal of Applied Poultry Research, 14*(2), 426-436.

Klein, E. D., Bhatt, R. S., & Zentall, T. R. (2005). Contrast and the justification of effort. *Psychonomic Bulletin & Review, 12*(2), 335-339.

Kliger, D., & Levy, O. (2003). Mood-induced variation in risk preferences. *Journal of Economic Behavior & Organization, 52*(4), 573-584.

Korsten, P., Vedder, O., Szentirmai, I., & Komdeur, J. (2007). Absence of status signalling by structurally based ultraviolet plumage in wintering blue tits (Cyanistes caeruleus). *Behavioral Ecology and Sociobiology, 61*(12), 1933-1943.

Kreft, I., & De Leeuw, J. (1998). *Introducing Multilevel Modeling*. London, U.K.: Sage Publications.

Krohn, T. C., Hansen, A. K., & Dragsted, N. (2003). Telemetry as a method for measuring the impact of housing conditions on rats' welfare. *Animal Welfare, 12*(1), 53-62.

Lane, R. D., Ahern, G. L., Schwartz, G. E., & Kaszniak, A. W. (1997). Is alexithymia the emotional equivalent of blindsight? *Biological Psychiatry, 42*(9), 834-844.

Langer, J., Wapner, S., & Werner, H. (1961). Effect of Danger Upon Experience of Time. *American Journal of Psychology, 74*(1), 94-&.

Latham, N. R., & Mason, G. (2006). *We've got to get out of this place: frustration, enrichment and the development of stereotypies in laboratory mice (Mus musculus)*. Paper presented at the International Congress of the ISAE, Bristol, UK.

Lavender, A., & Watkins, E. (2004). Rumination and future thinking in depression. *British Journal of Clinical Psychology, 43*, 129-142.

Lavie, N., Hirst, A., de Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology-General, 133*(3), 339-354.

Lawrence, P. A. (2007). The mismeasurement of science. *Current Biology, 17*(15), R583-R585.

Lawson, C., MacLeod, C., & Hammond, G. (2002). Interpretation revealed in the blink of an eye: Depressive bias in the resolution of ambiguity. *Journal of Abnormal Psychology, 111*(2), 321-328.

LeDoux, J. (1996). *The emotional brain: the mysterious underpinnings of emotional life*. New York: Simon & Schuster.

Lench, H. C., & Ditto, P. H. (2008). Automatic optimism: Biased use of base rate information for positive and negative events. *Journal of Experimental Social Psychology, 44*(3), 631-639.

Leone, E. H., Estevez, I., & Christman, M. C. (2007). Environmental complexity and group size: Immediate effects on use of space by domestic fowl. *Applied Animal Behaviour Science, 102*(1-2), 39-52.

Levine, J. (1983). Materialism and qualia: the explanatory gap. *Pacific Philosophical Quarterly, 64*, 354-361.

Libet, B. (1985). Unconscious Cerebral Initiative and the Role of Conscious Will in Voluntary Action. *Behavioral and Brain Sciences, 8*(4), 529-539.

Lindstrom, L., Rowe, C., & Guilford, T. (2001). Pyrazine odour makes visually conspicuous prey aversive. *Proceedings of the Royal Society of London Series B-Biological Sciences, 268*(1463), 159-162.

Loewenstein, G. F. (1996). Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes,* 65(3), 272-292.

Loewenstein, G. F., & Lerner, J. S. (2003). The Role of Affect in Decision Making. In R. J. Davidson, K. R. Scherer & H. H. Goldsmith (Eds.), *Handbook of Affective Sciences*. Oxford: Oxford University Press.

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin, 127*(2), 267-286.

Logan, F. A. (1965). Decision Making by Rats - Delay Versus Amount of Reward. *Journal of Comparative and Physiological Psychology, 59*(1), 1-&.

Lyubomirsky, S., & Nolenhoeksema, S. (1995). Effects of Self-Focused Rumination on Negative Thinking and Interpersonal Problem-Solving. *Journal of Personality and Social Psychology, 69*(1), 176-190.

MacLeod, A. K. (1999). Prospective cognitions. In T. Dalgleish & M. J. Power (Eds.), *Handbook of cognition and emotion* (pp. 267-280). Chichester: John Wiley.

MacLeod, A. K., Byrne, A., & Valentine, J. D. (1996). Affect, Emotional Disorder, and Future-directed Thinking. *Cognition & Emotion, 10*(1), 69-86.

MacLeod, A. K., Tata, P., Kentish, J., & Jacobsen, H. (1997). Retrospective and prospective cognitions in anxiety and depression. *Cognition & Emotion, 11*(4), 467-479.

MacLeod, C. (1999). Anxiety and Anxiety Disorders. In T. Dagleish & M. Power (Eds.), *Handbook of Cognition and Emotion*. Chichester: John Wiley & Sons Ltd.

Macleod, C., & Cohen, I. L. (1993). Anxiety and the Interpretation of Ambiguity - a Text Comprehension Study. *Journal of Abnormal Psychology, 102*(2), 238-247.

Macleod, C., Mathews, A., & Tata, P. (1986). Attentional Bias in Emotional Disorders. *Journal of Abnormal Psychology, 95*(1), 15-20.

Magrath, M. J. L., Brouwer, L., & Komdeur, J. (2003). Egg size and laying order in relation to offspring sex in the extreme sexually size dimorphic brown songlark, Cinclorhamphus cruralis. *Behavioral Ecology and Sociobiology, 54*(3), 240-248.

Maner, J. K., Richey, J. A., Cromer, K., Mallott, M., Lejuez, C. W., Joiner, T. E., et al. (2007). Dispositional anxiety and risk-avoidant decision-making. *Personality and Individual Differences, 42*(4), 665-675.

Maner, J. K., & Schmidt, N. B. (2006). The role of risk avoidance in anxiety. *Behavior Therapy, 37*(2), 181-189.

Marin, M. T., Cruz, F. C., & Planeta, C. S. (2007). Chronic restraint or variable stresses differently affect the behavior, corticosterone secretion and body weight in rats. *Physiology & Behavior, 90*(1), 29-35.

Marples, N. M., & Kelly, D. J. (1999). Neophobia and dietary Conservatism: Two distinct processes? *Evolutionary Ecology, 13*(7-8), 641-653.

Marples, N. M., & Roper, T. J. (1996). Effects of novel colour and smell on the response of naive chicks towards food and water. *Animal Behaviour, 51*, 1417-1424.

Marti, O., Marti, J., & Armario, A. (1994). Effects of Chronic Stress on Food-Intake in Rats - Influence of Stressor Intensity and Duration of Daily Exposure. *Physiology & Behavior, 55*(4), 747-753.

Martin, P., & Bateson, P. (1993). *Measuring Behaviour.* Cambridge, UK: Cambridge University Press.

Marx, G., Leppelt, J., & Ellendorff, F. (2001). Vocalisation in chicks (Gallus gallus dom.) during stepwise social isolation. *Applied Animal Behaviour Science, 75*(1), 61-74.

Maslova, L. N., Bulygina, V. V., & Markel, A. L. (2002). Chronic stress during prepubertal development: immediate and long-lasting effects on arterial blood pressure and anxiety-related behavior. *Psychoneuroendocrinology, 27*(5), 549-561.

Mason, G., & Mendl, M. (1993). Why is there no simple way of measuring animal welfare? *Animal Welfare, 2*, 301-319.

Matell, M. S., & Meck, W. H. (1999). Reinforcement-induced within-trial resetting of an internal clock. *Behavioural Processes, 45*(1-3), 159-171.

Matheson, S. M., Asher, L., & Bateson, M. (2008). Larger, enriched cages are associated with 'optimistic' response biases in captive European starlings (Sturnus vulgaris). *Applied Animal Behaviour Science, 109*(2-4), 374-383.

Mathews, A., & Macleod, C. (1985). Selective Processing of Threat Cues in Anxiety-States. *Behaviour Research and Therapy, 23*(5), 563-569.

Mathews, A., Richards, A., & Eysenck, M. (1989). Interpretation of homophones related to threat in anxiety states. *Journal of Abnormal Psychology, 98*(1), 31-34.

Matuszewich, L., Karney, J. J., Carter, S. R., Janasik, S. P., O'Brien, J. L., & Friedman, R. D. (2007). The delayed effects of chronic unpredictable stress on anxiety measures. *Physiology & Behavior, 90*(4), 674-681.

Meehan, C. L., & Mench, J. A. (2007). The challenge of challenge: Can problem solving opportunities enhance animal welfare? *Applied Animal Behaviour Science, 102*(3-4), 246-261.

Meier, B. P., & Robinson, M. D. (2006). Does "feeling down" mean seeing down? Depressive symptoms and vertical selective attention. *Journal of Research in Personality, 40*(4), 451-461.

Meier, B. P., Robinson, M. D., Crawford, L. E., & Ahlvers, W. J. (2007). When "light" and "dark" thoughts become light and dark responses: Affect biases brightness judgments. *Emotion, 7*(2), 366-376.

Mellers, B., Schwartz, A., & Ritov, I. (1999). Emotion-based choice. *Journal of Experimental Psychology-General, 128*(3), 332-345.

Menargues, A., Urios, V., & Mauri, M. (2008). Welfare assessment of captive Asian elephants (Elephas maximus) and Indian rhinoceros (Rhinoceros unicornis) using salivary cortisol measurement. *Animal Welfare, 17*(3), 305-312.

Mendl, M. (1999). Performing under pressure: stress and cognitive function. *Applied Animal Behaviour Science, 65*(3), 221-244.

Mendl, M., & Paul, E. S. (2004). Consciousness, emotion and animal welfare: insights from cognitive science. *Animal Welfare, 13*, S17-S25.

Mendl, M., & Paul, E. S. (2008). Do animals live in the present? Current evidence and implications for welfare. *Applied Animal Behaviour Science, 113*(4), 357-382.

Mialet, J. P., Pope, H. G., & YurgelunTodd, D. (1996). Impaired attention in depressive states: A non-specific deficit? *Psychological Medicine, 26*(5), 1009-1020.

Miklosi, A., Gonda, Z., Osorio, D., & Farzin, A. (2002). The effects of the visual environment on responses to colour by domestic chicks. *Journal of Comparative Physiology a-Neuroethology Sensory Neural and Behavioral Physiology, 188*(2), 135-140.

Milad, M. R., Wright, C. I., Orr, S. P., Pitman, R. K., Quirk, G. J., & Rauch, S. L. (2007). Recall of fear extinction in humans activates the ventromedial prefrontal cortex and hippocampus in concert. *Biological Psychiatry, 62*(5), 446-454.

Miles, H., MacLeod, A. K., & Pote, H. (2004). Retrospective and prospective cognitions in adolescents: anxiety, depression, and positive and negative affect. *Journal of Adolescence, 27*(6), 691-701.

Mineka, S., Rafaeli, E., & Yovel, I. (2003). Cognitive Biases in Emotional Disorders: Information Processing and Social-Cognitive Perspectives. In R. J. Davidson, K. R. Scherer & H. H. Goldsmith (Eds.), *Handbook of Affective Sciences*. Oxford: Oxford University Press.

Mineka, S., Watson, D., & Clark, L. A. (1998). Comorbidity of anxiety and unipolar mood disorders. *Annual Review of Psychology, 49*, 377-412.

Mineur, Y. S., Belzung, C., & Crusio, W. E. (2006). Effects of unpredictable chronic mild stress on anxiety and depression-like behavior in mice. *Behavioural Brain Research, 175*(1), 43-50.

Mitra, R., Vyas, A., Chatterjee, G., & Chattarji, S. (2005). Chronic-stress induced modulation of different states of anxiety-like behavior in female rats. *Neuroscience Letters, 383*(3), 278-283.

Mitte, K. (2007). Anxiety and risky decision-making: The role of subjective probability and subjective costs of negative events. *Personality and Individual Differences, 43*(2), 243-253.

Moberg, G. P., & Mench, J. A. (2000). *The Biology of Animal Stress*. Wallingford, Oxon.: CABI Publishing.

Mogg, K., Bradbury, K. E., & Bradley, B. P. (2006). Interpretation of ambiguous information in clinical depression. *Behaviour Research and Therapy, 44*(10), 1411-1419.

Mogg, K., & Bradley, B. P. (2005). Attentional bias in generalized anxiety disorder versus depressive disorder. *Cognitive Therapy and Research, 29*(1), 29-45.

Mogg, K., Mathews, A., & Eysenck, M. (1992). Attentional Bias to Threat in Clinical Anxiety-States. *Cognition & Emotion, 6*(2), 149-159.

Mogg, K., Mathews, A., & Weinman, J. (1989). Selective Processing of Threat Cues in Anxiety-States - a Replication. *Behaviour Research and Therapy, 27*(4), 317-323.

Moody, T. C. (1994). Conversations with Zombies. *Journal of Consciousness Studies, 1*(2), 196-200.

Moore, B. C. J. (2004). *An Introduction to the Psychology of Hearing*. London: Academic Press.

Moore, M. T., & Fresco, D. M. (2007). Depressive realism and attributional style: Implications for individuals at risk for depression. *Behavior Therapy, 38*(2), 144-154.

Morato, S., & Brandao, M. L. (1997). Paradoxical increase of exploratory behavior in the elevated plus-maze by rats exposed to two kinds of aversive stimuli. *Brazilian Journal of Medical and Biological Research, 30*(9), 1113-1120.

Morgan, K. N., & Tromborg, C. T. (2007). Sources of stress in captivity. *Applied Animal Behaviour Science, 102*(3-4), 262-302.

Muller, W., Dijkstra, C., & Groothuis, T. G. G. (2003). Inter-sexual differences in T-cell-mediated immunity of black-headed gull chicks (Larus ridibundus) depend on the hatching order. *Behavioral Ecology and Sociobiology, 55*(1), 80-86.

Mundt, C., Richter, P., van Hees, H., & Stumpf, T. (1998). Time-experience and time-estimation in depressive patients. *Nervenarzt, 69*(1), 38-45.

Murphy, F. C., Rubinsztein, J. S., Michael, A., Rogers, R. D., Robbins, T. W., Paykel, E. S., et al. (2001). Decision-making cognition in mania and depression. *Psychological Medicine, 31*(4), 679-693.

Myrtek, M., & Brugner, G. (1996). Perception of emotions in everyday life: Studies with patients and normals. *Biological Psychology, 42*(1-2), 147-164.

Nader, K., & Balleine, B. (2007). Ambiguity and anxiety: when a glass half full is empty. *Nature Neuroscience, 10*(7), 807-808.

Nagel, T. (1974). What Is It Like To Be A Bat. *Philosophical Review, 83*(4), 435-450.

Nakagawa, S., & Cuthill, I. C. (2007). Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological Reviews, 82*(4), 591-605.

Nettle, D. (2004). Evolutionary origins of depression: a review and reformulation. *Journal Of Affective Disorders, 81*(2), 91-102.

Newberry, R. C., & Shackleton, D. M. (1997). Use of visual cover by domestic fowl: A Venetian blind effect? *Animal Behaviour, 54*, 387-395.

Nezlek, J. B. (2001). Multilevel random coefficient analyses of event- and interval-contingent data in social and personality psychology research. *Personality and Social Psychology Bulletin, 27*(7), 771-785.

Nicol, C. J. (1992). Effects of Environmental Enrichment and Gentle Handling on Behavior and Fear Responses of Transported Broilers. *Applied Animal Behaviour Science, 33*(4), 367-380.

Norušis, M. J. (1999). *SPSS Regression Models 10.0*. Chicago: SPSS Inc.

Nygren, T. E., Isen, A. M., Taylor, P. J., & Dulin, J. (1996). The Influence of Positive Affect on the Decision Rule in Risk Situations: Focus on Outcome (and Especially Avoidance of Loss) Rather Than Probability. *Organizational Behavior and Human Decision Processes, 66*(1), 59-72.

O'Connor, R., & Eikelboom, R. (2000). The effects of changes in housing on feeding and wheel running. *Physiology & Behavior, 68*(3), 361-371.

Ochsner, K. N., & Phelps, E. (2007). Emerging perspectives on emotion-cognition interactions. *Trends in Cognitive Sciences, 11*(8), 317-318.

Ohl, F. (2003). Testing for anxiety. *Clinical Neuroscience Research, 3*(4-5), 233-238.

Ohl, F., Toschi, N., Wigger, A., Henniger, M. S. H., & Landgraf, R. (2001). Dimensions of emotionality in a rat model of innate anxiety. *Behavioral Neuroscience, 115*(2), 429-436.

Olgiati, P., Serretti, A., & Colombo, C. (2006). Retrospective analysis of psychomotor agitation, hypomanic symptoms, and suicidal ideation in unipolar depression. *Depression and Anxiety, 23*(7), 389-397.

Osorio, D., Miklosi, A., & Gonda, Z. (1999). Visual ecology and perception of coloration patterns by domestic chicks. *Evolutionary Ecology, 13*(7-8), 673-689.

Panayiotou, G. (2008). Emotional dimensions reflected in ratings of affective scripts. *Personality and Individual Differences, 44*(8), 1795-1806.

Panksepp, J. (2003). At the interface of the affective, behavioral, and cognitive neurosciences: Decoding the emotional feelings of the brain. *Brain And Cognition, 52*(1), 4-14.

Partan, S. R., & Marler, P. (2005). Issues in the classification of multimodal communication signals. *American Naturalist, 166*(2), 231-245.

Paul, E. S., Harding, E. J., & Mendl, M. (2005). Measuring emotional processes in animals: the utility of a cognitive approach. *Neuroscience & Biobehavioral Reviews, 29*(3), 469-491.

Paul, E. S., & Mendl, M. (in prep.). *Is it possible to measure subjective emotion in animals?*Unpublished manuscript.

Paulignan, Y., Mackenzie, C., Marteniuk, R., & Jeannerod, M. (1990). The Coupling of Arm and Finger Movements During Prehension. *Experimental Brain Research, 79*(2), 431-435.

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience, 9*(2), 148-158.

Pfister, H. P. (1979). The glucocorticosterone response to novelty as a psychological stressor. *Physiology & Behavior, 23*(4), 649-652.

Phelps, E. A. (2006). Emotion and cognition: Insights from studies of the human amygdala. *Annual Review of Psychology, 57*, 27-53.

Phillips, A. G., & Barr, A. M. (1997). Effects of chronic mild stress on motivation for sucrose: mixed messages. *Psychopharmacology, 134*(4), 361-362.

Pier, M., Hulstijn, W., & Sabbe, B. G. C. (2004). Differential patterns of psychomotor functioning in unmedicated melancholic and nonmelancholic depressed patients. *Journal of Psychiatric Research, 38*(4), 425-435.

Pinker, S. (1994). *The Language Instinct*. London: Penguin.

Popovic, M., Caballero-Bleda, M., Puelles, L., & Guerri, C. (2004). Multiple binge alcohol consumption during rat adolescence increases anxiety but does not impair retention in the passive avoidance task. *Neuroscience Letters, 357*(2), 79-82.

Powell, S. B., Geyer, M. A., Gallagher, D., & Paulus, M. P. (2004). The balance between approach and avoidance behaviors in a novel object exploration paradigm in mice. *Behavioural Brain Research, 152*(2), 341-349.

Power, M. (1999). Sadness and Its Disorders. In T. Dalgleish & M. Power (Eds.), *Handbook of Cognition and Emotion*. Chichester: John Wiley & Sons Ltd.

Prut, L., & Belzung, C. (2003). The open field as a paradigm to measure the effects of drugs on anxiety-like behaviours: a review. *European Journal of Pharmacology, 463*, 3-33.

Quené, H., & van den Bergh, H. (2004). On multi-level modeling of data from repeated measures designs: a tutorial. *Speech Communication, 43*, 103-121.

Quinn, G. P., & Keough, M. J. (2002). *Experimental Design and Data Analysis for Biologists*. Cambridge: Cambridge University Press.

Rasbash, J., Steele, F., Browne, W., & Prosser, B. (2005). *A User's Guide to MLwiN Version 2.0*. University of Bristol: Centre for Multilevel Modelling.

Reed, P. (1991). Multiple Determinants of the Effects of Reinforcement Magnitude on Free-Operant Response Rates. *Journal of the Experimental Analysis of Behavior, 55*(1), 109-123.

Regolin, L., Vallortigara, G., & Zanforlin, M. (1995). Detour Behavior in the Domestic Chick - Searching for a Disappearing Prey or a Disappearing Social Partner. *Animal Behaviour, 50*, 203-211.

Reid, I., Forbes, N., Stewart, C., & Matthews, K. (1997). Chronic mild stress and depressive disorder: a useful new model? *Psychopharmacology, 134*(4), 365-367.

Richards, A., & French, C. C. (1992). An Anxiety-Related Bias in Semantic Activation When Processing Threat/Neutral Homographs. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 45*(3), 503-525.

Richards, A., French, C. C., Calder, A. J., Webb, B., Fox, R., & Young, A. W. (2002). Anxiety-Related Bias in the Classification of Emotionally Ambiguous Facial Expressions. *Emotion, 2*(3), 273-287.

Richards, J. B., Mitchell, S. H., DeWit, H., & Seiden, L. S. (1997). Determination of discount functions in rats with an adjusting-amount procedure. *Journal of the Experimental Analysis of Behavior, 67*(3), 353-366.

Rodgers, R. J. (1997). Animal models of 'anxiety': where next? *Behavioural Pharmacology, 8*(6-7), 477-496.

Rogers, L. J. (1995). *The Development of Brain and Behaviour in the Chicken*. Wallingford, Oxon.: CAB International.

Rolls, E. T. (1999). *The Brain and Emotion*. Oxford, UK: Oxford University Press.

Rolls, E. T. (2005). *Emotion Explained*. Oxford: Oxford University Press.

Rose, E. J., & Ebmeier, K. P. (2006). Pattern of impaired working memory during major depression. *Journal of Affective Disorders, 90*(2-3), 149-161.

Roth, K. A., & Katz, R. J. (1979). Stress, Behavioral Arousal, and Open-Field Activity - Reexamination of Emotionality in the Rat. *Neuroscience and Biobehavioral Reviews, 3*(4), 247-263.

Rowe, C. (1999). Receiver psychology and the evolution of multicomponent signals. *Animal Behaviour, 58*, 921-931.

Rowe, C., & Guilford, T. (1996). Hidden colour aversions in domestic clicks triggered by pyrazine odours of insect warning displays. *Nature, 383*(6600), 520-522.

Rowe, C., & Guilford, T. (1999a). The evolution of multimodal warning displays. *Evolutionary Ecology, 13*(7-8), 655-671.

Rowe, C., & Guilford, T. (1999b). Novelty effects in a multimodal warning signal. *Animal Behaviour, 57*, 341-346.

Rowe, C., & Skelhorn, J. (2005). Colour biases are a question of taste. *Animal Behaviour, 69*, 587-594.

Roy, V., & Chapillon, P. (2004). Further evidences that risk assessment and object exploration behaviours are useful to evaluate emotional reactivity in rodents. *Behavioural Brain Research, 154*(2), 439-448.

Rubinsztein, J. S., Michael, A., Underwood, B. R., Tempest, M., & Sahakian, B. J. (2006). Impaired cognition and decision-making in bipolar depression but no 'affective bias' evident. *Psychological Medicine, 36*(5), 629-639.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*, 1161-1178.

Russo, R., Fox, E., & Bowles, R. J. (1999). On the status of implicit memory bias in anxiety. *Cognition & Emotion, 13*(4), 435-456.

Ruxton, G. D., Sherratt, T. N., & Speed, M. P. (2004). *Avoiding Attack: The Evolutionary Ecology of Crypsis, Warning Signals and Mimicry*. Oxford: Oxford University Press.

Sandem, A. I., Janczak, A. M., Salte, R., & Brotastad, B. O. (2006). The use of diazepam as a pharmacological validation of eye white as an indicator of emotional state in dairy cows. *Applied Animal Behaviour Science, 96*(3-4), 177-183.

Sanotra, G. S., Vestergaard, K. S., Agger, J. F., & Lawson, L. G. (1995). The Relative Preferences for Feathers, Straw, Wood-Shavings and Sand for Dustbathing, Pecking and Scratching in Domestic Chicks. *Applied Animal Behaviour Science, 43*(4), 263-277.

Santos, J., Benjamin, M., Yang, P. C., Prior, T., & Perdue, M. H. (2000). Chronic stress impairs rat growth and jejunal epithelial barrier function: role of mast cells.

American Journal of Physiology-Gastrointestinal and Liver Physiology, 278(6), G847-G854.

Schaefer, H. M., Levey, D. J., Schaefer, V., & Avery, M. L. (2006). The role of chromatic and achromatic signals for fruit detection by birds. *Behavioral Ecology, 17*(5), 784-789.

Scherer, K. R. (1999). Appraisal Theory. In T. Dagleish & M. Power (Eds.), *Handbook of Cognition and Emotion* (pp. 637-663). Chichester, England: John Wiley & Sons Ltd.

Scherer, K. R. (2001). Appraisal considered as a process of multi-level sequential checking. In K. R. Scherer, A. Schorr & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, methods, research.* New York: Oxford University Press.

Scherer, K. R. (2003). Introduction: Cognitive Components of Emotion. In R. J. Davidson, K. R. Scherer & H. H. Goldsmith (Eds.), *Handbook of Affective Sciences.* Oxford: Oxford University Press.

Schmidt, V., Schaefer, H. M., & Winkler, H. (2004). Conspicuousness, not colour as foraging cue in plant-animal signalling. *Oikos, 106*(3), 551-557.

Schrijvers, D., Hulstijn, W., & Sabbe, B. G. C. (2008). Psychomotor symptoms in depression: A diagnostic, pathophysiological and therapeutic tool. *Journal of Affective Disorders, 109*(1-2), 1-20.

Schutz, K., Davison, D., & Matthews, L. (2006). Do different levels of moderate feed deprivation in dairy cows affect feeding motivation? *Applied Animal Behaviour Science, 101*(3-4), 253-263.

Schwarz, N., & Clore, G. L. (1983). Mood, Misattribution, and Judgments of Well-Being - Informative and Directive Functions of Affective States. *Journal of Personality and Social Psychology, 45*(3), 513-523.

Seth, A. K., Baars, B. J., & Edelman, D. B. (2005). Criteria for consciousness in humans and other mammals. *Consciousness and Cognition, 14*(1), 119-139.

Shettleworth, S. J. (1998). *Cognition, Evolution, and Behavior.* Oxford, UK: Oxford University Press.

Shrauger, J. S., Mariano, E., & Walter, T. J. (1998). Depressive symptoms and accuracy in the prediction of future events. *Personality and Social Psychology Bulletin, 24*(8), 880-892.

Simon, P., Dupuis, R., & Costentin, J. (1994). Thigmotaxis as an Index of Anxiety in Mice - Influence of Dopaminergic Transmissions. *Behavioural Brain Research, 61*(1), 59-64.

Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2004). Risk as analysis and risk as feelings: Some thoughts about affect, reason, risk, and rationality. *Risk Analysis, 24*(2), 311-322.

Smith, J. D., & Washburn, D. A. (2005). Uncertainty monitoring and metacognition by animals. *Current Directions in Psychological Science, 14*(1), 19-24.

Sneddon, L. U., Braithwaite, V. A., & Gentle, M. J. (2003). Novel object test: Examining nociception and fear in the rainbow trout. *Journal of Pain, 4*(8), 431-440.

Snowden, R., Thompson, P., & Troscianko, T. (2006). *Basic Vision: An Introduction to Visual Perception*. Oxford: Oxford University Press.

Sokal, R. R., & Rohlf, F. J. (1995). *Biometry*. New York: W.H. Freeman & Co.

Steinberg, H., & Watson, R. H. J. (1960). Failure of Growth in Disturbed Laboratory Rats. *Nature, 185*, 615-616.

Stober, J. (2000). Prospective cognitions in anxiety and depression: Replication and methodological extension. *Cognition & Emotion, 14*(5), 725-729.

Stoddart, S. D. R., Craddock, N. J., & Jones, L. A. (2007). Differentiation of executive and attention impairments in affective illness. *Psychological Medicine, 37*(11), 1613-1623.

Stone, L. A., & Nielson, K. A. (2001). Intact physiological responses to arousal with impaired emotional recognition in alexithymia. *Psychotherapy and Psychosomatics, 70*(2), 92-102.

Strekalova, T., Spanagel, R., Dolgov, O., & Bartsch, D. (2005). Stress-induced hyperlocomotion as a confounding factor in anxiety and depression models in mice. *Behavioural Pharmacology, 16*(3), 171-180.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643-662.

Suarez, S. D., & Gallup, G. G. (1983). Social Reinstatement and Open-Field Testing in Chickens. *Animal Learning & Behavior, 11*(1), 119-126.

Sufka, K. J., Feltenstein, M. W., Warnick, J. E., Acevedo, E. O., Webb, H. E., & Cartwright, C. M. (2006). Modeling the anxiety-depression continuum hypothesis in domestic fowl chicks. *Behavioural Pharmacology, 17*(8), 681-689.

Syka, J., Rybalko, N., Brozek, G., & Jilek, M. (1996). Auditory frequency and intensity discrimination in pigmented rats. *Hearing Research, 100*(1-2), 107-113.

Talwar, S. K., & Gerstein, G. L. (1998). Auditory frequency discrimination in the white rat. *Hearing Research, 126*(1-2), 135-150.

Tamir, M., & Robinson, M. D. (2007). The happy spotlight: Positive mood and selective attention to rewarding information. *Personality and Social Psychology Bulletin, 33*(8), 1124-1136.

Taylor, S. E., & Brown, J. D. (1988). Illusion and Well-Being - a Social Psychological Perspective on Mental-Health. *Psychological Bulletin, 103*(2), 193-210.

Teachman, B. A., Stefanucci, J. K., Clerkin, E. M., Cody, M. W., & Proffitt, D. R. (2008). A new mode of fear expression: Perceptual bias in height fear. *Emotion, 8*(2), 296-301.

Teasdale, J. D., & Fogarty, S. J. (1979). Differential Effects of Induced Mood on Retrieval of Pleasant and Unpleasant Events from Episodic Memory. *Journal of Abnormal Psychology, 88*(3), 248-257.

Tellegen, A. (1985). Structures of mood and personality and their relevance to assessing anxiety with an emphasis on self-report. In A. H. Tuma & J. D. Maser (Eds.), *Anxiety and the anxiety disorders*. Hillsdale, NJ.: Lawrence Erlbaum Associates Inc.

Tomarken, A. J., Cook, M., & Mineka, S. (1989). Fear-Relevant Selective Associations and Covariation Bias. *Journal of Abnormal Psychology, 98*(4), 381-394.

Tsetsenis, T., Ma, X. H., Lo Iacono, L., Beck, S. G., & Gross, C. (2007). Suppression of conditioning to ambiguous cues by pharmacogenetic inhibition of the dentate gyrus. *Nature Neuroscience, 10*(7), 896-902.

Tsuchiya, N., & Adolphs, R. (2007). Emotion and consciousness. *Trends in Cognitive Sciences, 11*(4), 158-167.

Tversky, A., & Kahneman, D. (1973). Availability - Heuristic for Judging Frequency and Probability. *Cognitive Psychology, 5*(2), 207-232.

Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social-emotional development. *Psychological Bulletin, 134*(3), 383-403.

Vallortigara, G., Cailotto, M., & Zanforlin, M. (1990). Sex-Differences in Social Reinstatement Motivation of the Domestic Chick (Gallus-Gallus) Revealed by Runway Tests with Social and Nonsocial Reinforcement. *Journal of Comparative Psychology, 104*(4), 361-367.

Van Damme, S., Crombez, G., Hermans, D., Koster, E. H. W., & Eccleston, C. (2006). The role of extinction and reinstatement in attentional bias to threat: A conditioning approach. *Behaviour Research and Therapy, 44*(11), 1555-1563.

van Gaalen, M. M., & Steckler, T. (2000). Behavioural analysis of four mouse strains in an anxiety test battery. *Behavioural Brain Research, 115*(1), 95-106.

van Praag, H., Kempermann, G., & Gage, F. H. (2000). Neural consequences of environmental enrichment. *Nature Reviews Neuroscience, 1*(3), 191-198.

Velmans, M. (2000). *Understanding Consciousness*. London and Philadelphia: Routledge.

VollmerConna, U., Wakefield, D., Lloyd, A., Hickie, I., Lemon, J., Bird, K. D., et al. (1997). Cognitive deficits in patients suffering from chronic fatigue syndrome, acute infective illness or depression. *British Journal of Psychiatry, 171*, 377-381.

von Neumann, J., & Morgenstern, O. (1947). *Theory of Games and Economic Behaviour*. Princeton, NJ: Princeton University Press.

Vyas, A., & Chattarji, S. (2004). Modulation of different states of anxiety-like behavior by chronic stress. *Behavioral Neuroscience, 118*(6), 1450-1454.

Walf, A. A., & Frye, C. A. (2007). The use of the elevated plus maze as an assay of anxiety-related behavior in rodents. *Nature Protocols, 2*(2), 322-328.

Warnick, J. E., Wicks, R. T., & Sufka, K. J. (2006). Modeling anxiety-like states: pharmacological characterization of the chick separation stress paradigm. *Behavioural Pharmacology, 17*(7), 581-587.

Watson, D., & Clark, L. A. (1984). Negative Affectivity - the Disposition to Experience Aversive Emotional States. *Psychological Bulletin, 96*(3), 465-490.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development And Validation Of Brief Measures Of Positive And Negative Affect - The Panas Scales. *Journal Of Personality And Social Psychology, 54*(6), 1063-1070.

Watson, D., Wiese, D., Vaidya, J., & Tellegen, A. (1999). The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence. *Journal Of Personality And Social Psychology, 76*(5), 820-838.

Watts, F. N., & Sharrock, R. (1984). Fear and Time-Estimation. *Perceptual and Motor Skills, 59*(2), 597-598.

Weeks, C. A., & Nicol, C. J. (2006). Behavioural needs, priorities and preferences of laying hens. *Worlds Poultry Science Journal, 62*(2), 296-307.

Wegner, D. M. (2003). The mind's best trick: how we experience conscious will. *Trends in Cognitive Sciences, 7*(2), 65-69.

Weiskrantz, L. (2001). Commentary responses and conscious awareness in humans: The implications for awareness in non-human animals. *Animal Welfare, 10*, S41-S46.

Wickens, T. D. (2002). *Elementary Signal Detection Theory*. Oxford: Oxford University Press.

Williams, J. M. G., Mathews, A., & MacLeod, C. (1996). The emotional stroop task and psychopathology. *Psychological Bulletin, 120*(1), 3-24.

Willner, P. (1997). Validity, reliability and utility of the chronic mild stress model of depression: a 10-year review and evaluation. *Psychopharmacology, 134*(4), 319-329.

Willner, P. (2005). Chronic mild stress (CMS) revisited: Consistency and behavioural-neurobiological concordance in the effects of CMS. *Neuropsychobiology, 52*(2), 90-110.

Willner, P., Muscat, R., & Mariusz, P. (1992). Chronic Mild Stress-Induced Anhedonia: A Realistic Animal Model of Depression. *Neuroscience and Biobehavioral Reviews, 16*, 525-534.

Willner, P., Towell, A., Sampson, D., Sophokleous, S., & Muscat, R. (1987). Reduction Of Sucrose Preference By Chronic Unpredictable Mild Stress, And Its Restoration By A Tricyclic Antidepressant. *Psychopharmacology, 93*(3), 358-364.

Willson, M. F., & Whelan, C. J. (1990). The Evolution of Fruit Color in Fleshy-Fruited Plants. *American Naturalist, 136*(6), 790-809.

Winkielman, P., Berridge, K. C., & Wilbarger, J. L. (2005). Unconscious Affective Reactions to Masked Happy Versus Angry Faces Influence Consumption Behavior and Judgments of Value. *Personality And Social Psychology Bulletin, 31*, 121-135.

Wittmann, M., & Paulus, M. P. (2008). Decision making, impulsivity and time perception. *Trends in Cognitive Sciences, 12*(1), 7-12.

Wurbel, H. (2001). Ideal homes? Housing effects on rodent brain and behaviour. *Trends in Neurosciences, 24*(4), 207-211.

Yalcin, I., Aksu, F., Bodard, S., Chalon, S., & Belzung, C. (2007). Antidepressant-like effect of tramadol in the unpredictable chronic mild stress procedure: possible involvement of the noradrenergic system. *Behavioural Pharmacology, 18*(7), 623-631.

Yoon, K. L., & Zinbarg, R. E. (2007). Threat is in the eye of the beholder: Social anxiety and the interpretation of ambiguous facial expressions. *Behaviour Research and Therapy, 45*(4), 839-847.

Yost, W. A. (2000). *Fundamentals of Hearing*. London: Academic Press.

Young, R. J. (2003). *Environmental Enrichment for Captive Animals*. Oxford, UK: Blackwell.

Yuen, K. S. L., & Lee, T. M. C. (2003). Could mood state affect risk-taking decisions? *Journal of Affective Disorders, 75*(1), 11-18.

Zar, J. H. (1996). *Biostatistical Analysis*. London, U.K.: Prentice Hall.

Zurita, A., Martijena, I., Cuadra, G., Brandao, M. L., & Molina, V. (2000). Early exposure to chronic variable stress facilitates the occurrence of anhedonia and enhanced emotional reactions to novel stressors: reversal by naltrexone pretreatment. *Behavioural Brain Research, 117*(1-2), 163-171.

UNIVERSITY OF BRISTOL LIBRARY

MEDICAL